Author:
Copeland, Conrad

Title:
Centres of Gravity

Gravity Models, Market Access, and the Drivers of Trade

Centres of Gravity:

Gravity Models, Market Access, and the Drivers of Trade

Conrad E. Copeland

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Social Sciences and Law.

School of Economics

February 2021

55,164 words

Abstract

In this dissertation, I analyse the drivers of trade, both modern and historical, and how these forces influence trade flows and growth outcomes. In the first chapter, I investigate the effects of ethnic linkages on trade between countries within Africa. I construct a newly digitised and geo-referenced dataset of historical trade routes for pre-colonial Africa. The empirical strategy exploits the role of historical trade routes in shaping the distribution of ethnic groups across countries to estimate the impact of shared ethnicity on modern trade flows. The impact of ethnic links is quantitatively similar to that of formal trade agreements, with trade more than doubling between countries that share a common ethnic group. The results are driven by a few ethnic groups creating powerful links between countries despite often being minority groups. In the second chapter, I estimate a modified gravity model that incorporates the impact of conflict in adjacent countries into transport costs. I use geographic variation in the location of countries and conflicts to identify an adjacent conflict effect that negatively impacts trade. Countries with already precarious access to external markets (such as landlocked countries) are more adversely affected by adjacent conflict, experiencing up to twice as severe a drop in trade as other countries. In the third chapter, I exploit an ancient textual source, the Periplus of the Erythraean Sea, to measure and model ancient trade in the Indian Ocean. I construct sailing times between ancient cities and estimate a gravity model: distance matters for trade in the region, but there are significant non-linear effects due to resource endowments. Trade appears to encourage economic growth in the long run, particularly for those cities that export a greater variety of products. Additionally, some cities appear to be constrained by an ancient version of the 'resource curse'.

Acknowledgements

I am extremely grateful to my supervisors, Helen Simpson and Yanos Zylberberg, for all their support. Their comments were always helpful and they tirelessly supported me in my research. I will always be thankful for their encouragement and inspiration to pursue unique ideas and avenues of inquiry.

I have been lucky to be a part of the School of Economics and the University of Bristol. I would like to thank all the staff, students, academics, and fellow PhD students who made my research experience more enjoyable and rewarding.

I would also like to acknowledge Carlos Llanos, David Harper, Cornelia McCarthy, and Irasema Alonso at New York University for all their help and support. I also thank Michael Vaney at the University of British Columbia for the frequent conversations and advice early in my academic career.

Finally, I would like to thank my partner, Katherine, for all her love, encouragement, and constant support. Thank you for keeping me grounded and your unflagging determination, even in moments when my own was waning. I could not have had a better companion on this journey.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

Table of Contents

List of Tables

List of Figures

Figure **Page**

Introduction

This dissertation is concerned with understanding the drivers of trade and the way in which these forces impact trade flows and growth outcomes. Social or cultural characteristics and geography can both influence the ability of countries (and cities) to access external markets in positive and negative ways. In the following chapters I explore how social elements can develop trade promoting networks, as well as how the geography of trade routes can both be a hazard to trade flows and facilitate them.

Market access and availability of trade networks are important factors determining how easily agents and countries can buy and sell goods internationally and limiting or encouraging these connections can have significant implications not just for trade but also wider economic outcomes. The three questions I examine are: 1) how ethnic networks influence aggregate trade patterns; 2) whether adjacent conflict negatively impacts trade flows; and 3) in what way geography and climate influenced long-distance trade networks in the late Iron Age Indian Ocean. The first question addresses how local networks can foster and develop trade. Whereas much research on social connections focuses on the micro-level outcomes, the networks created by these connections can have large aggregate effects. The second question seeks to answer how disruptions to infrastructure and access to external markets can impact trade in negative ways. Geography is often understood as a static feature of trade routes, but particular geographic circumstances can have important effects on these trade routes by exposing them to potential disruption and threatening trade flows through conflict. Finally, the third question asks how geography and climate can encourage the creation of ocean-spanning networks. The unique features of a geographic area, such as wind patterns, can facilitate trade by reducing the costs and risks of engaging in long-distance networks that would have otherwise been closed.

In the first chapter, I investigate how local ethnic networks influence trade within Africa. Commercial relationships require both trust and information and the potential connections between ethnic groups across national boundaries can foster economic linkages and facilitate trade by providing a basis for trust and the exchange of information. I develop a new dataset of ethnicity in Africa from never before used sources in order to test the impact of cross-border ethnic commonality on trade. The empirical approach exploits the exogenous nature of pre-colonial trade posts and routes across the continent to instrument for the modern ethnic link measure. These pre-colonial trade networks influenced pre-modern population movements and the intensity of countries' integration into these routes is correlated with modern shared ethnicities. I find that ethnic networks are strong promoters of trade, with ethnic links between countries more than doubling trade. The magnitude of the effect is on par with the impact of formal trade agreements between governments. The number of shared groups and the similarity in ethnic composition of countries tend to have less of an effect on trade, implying that the effect is driven by a few ethnic communities creating powerful links between countries despite often being minority groups.

The second chapter is in this same line of research, but looks at the potential negative effects of conflict through the disruption of trade routes. In it, I explore the effect of conflict in adjacent countries on trade. Conflict in adjacent countries can potentially block or disrupt vital trade routes, I therefore develop a modified gravity model to analyse the trade implications of adjacent conflict. Geography plays an important role in identifying the effect, with landlocked countries that rely on foreign ports to trade more at risk of disruption from adjacent conflict blocking access to those ports. I use this difference in geography to disentangle the different effects of adjacent conflict, specifically whether a country is landlocked or not and whether the adjacent conflict occurs in a shoreward country or an inland country. While trade on the whole is negatively impacted, I find that landlocked countries experience significantly more negative effects than coastal countries. Similarly, conflicts that occur in shoreward countries are more disruptive to trade than conflicts that occur inland. This implies a market access effect in that conflict in an adjacent country can disrupt trade routes, lowering trade. This effect is most pronounced for landlocked countries which need to ship goods through another country in order to access external markets, but the effect on coastal countries also indicates a broader, more disruptive effect across all types of trade routes—hindering trade for all neighbours.

For the third chapter, I take a longer view of history and analyse the nature of trade networks in the first half of the first millennium. To do this, I build an entirely new dataset of trade goods and partners from ancient textual sources and use a Ricardian gravity model to analyse the impact of distance on trade flows in the Indian Ocean during this period. The unique seasonal wind patterns of the monsoon allowed maritime trade between distant cities, from the east coast of Africa to India. Trade in the Indian Ocean in the Iron Age was significantly influenced by distance, but even more so by travel time: sailing time between ports is a better fit to the model than simple distance indicating that wind currents and the conditions presented by the monsoon system created a sustainable network of long-distance trade. Further, the impact of distance is non-linear, with ports that are close together breaking from the standard gravity impact of distance and trading less. I find that this appears to be related to an endowment effect, with cities that have similar baskets of export goods trading less than those that are more different. The geographic nature of resource endowments means that cities that are close together tend to share common resource endowments, producing this non-linearity in the relationship between distance and trade.

I further find that trade appears to encourage economic growth in the long-run, with cities that export a greater variety of products having significantly more activity in their vicinities in the following 500 years. This is most pronounced for cities that export manufactured products and other 'high-value' goods, while those that are more dependent on raw materials and cash-crops such as spices and incense faring worse—implying an ancient version of the 'resource curse'.

The insights of this work provide perspective on how market access and trade networks can grow and develop. Social connections can build trust and reciprocity that form the foundation of trade networks between communities in different countries. Geography can expose trade routes to dangers that create risk for trade; tenuous connections to foreign ports can be easily threatened by conflict related disruptions. However, geography can also create ideal circumstances for facilitating long-distance trade networks; historically, climate and wind patterns could dictate the ability of what would otherwise have been distant cities to trade with each other in robust and sophisticated ways. These features are what influence trade flows and networks, and understanding them will help us to comprehend how and why trade develops between some places and not others.

Chapter I

Bridging New Divides: Ethnic Linkages and Trade in Africa

1 Introduction

The question of what determines trade between countries in Africa has tended to focus on factors from more recent history such as the end of colonialism (Head et al., 2010; Michalopoulos and Papaioannou, 2016; Mitchener and Weidenmier, 2008) or the formal economic linkages between governments on the continent (Geda and Kebret, 2007; Glick and Rose, 2002; Rose, 2004). Less attention has been given to informal institutional connections between countries, such as networks based on ethnicity. Through the processes of colonisation and decolonisation, groups that had been historically connected were split between various countries following arbitrary colonial boundaries. As a result, ethnic groups located along historical trade routes were often distributed among multiple states and sometimes quite distant from one another. Commercial relationships are hindered by the presence of moral hazard and information asymmetries; and the potential connections between these groups across national boundaries may foster economic linkages or facilitate trade by providing a basis for trust and information exchange.

This chapter examines the impact of ethnic networks on trade between countries in Africa. To do this, I construct a new dataset of ethnic groups and common ethnicities for the continent at a level of detail not previously available. The dataset uses local missionary data for ethnic groups that measure both ethnic population size and geographic location. The empirical strategy exploits the role of historical trade routes in shaping the distribution of ethnic groups across countries, a variation which is exogenous to later migration and population flows, and the fragmentation of Africa into different post-colonial states. I find that ethnic links increase trade between countries: trade more than doubles between countries that share a common ethnic community. This effect is similar to that of customs unions and regional trade agreements and is in line with, but distinct from, recent work on linguistic links.[1]

[1] Melitz (2008); Melitz and Toubal (2014); Selmier II and Oh (2013).

Importantly, this effect is driven by a few ethnic communities creating powerful links between countries despite often being minority groups. The results indicate that the extensive margin is driving the results: having an ethnic link is more important than the number of links or population sizes. Minority ethnic groups, particularly small enclaves within countries, are developing and exploiting networks across borders to facilitate trade. In this way, these networks are addressing issues of trust and information not covered by formal institutional systems.

This chapter contributes to the literature by combining ethnic population data with historical data on trade routes to investigate the effect of common ethnic links at a continent-wide scale and on national-level trade flows. This contribution is closest to Melitz and Toubal (2014) which deals with common spoken language (rather than common official language at the state level) and Aker et al. (2014) which looks at cross-border ethnic effects on trade between markets in Niger and Nigeria. This chapter expands on Melitz and Toubal (2014) by looking specifically at ethnicity which may differ from linguistic identities; it also incorporates historical trade routes and the impact they had on the geographic spread of ethnic groups to better identify the effect. The focus of this work on the effect of ethnic ties at a continent-wide level and in terms of total national trade flows compliments the emphasis of Aker et al. (2014) on micro-level effects. By creating this synthesis of ethnic links and national trade flows, this chapter presents evidence for the scalability of network mechanisms to influence national trade flows and the role of ethnic networks specifically in this process.

An important issue in estimating the impact of ethnic links on trade is reverse causality and omitted variables. It could be that trade directly influences the migration of ethnic groups and this could bias results. Other unobserved linkages might also exist between trade partners, such as similarity in consumption preferences, which could correlate with similarities in ethnic populations. The empirical strategy therefore exploits pre-colonial trade routes in order to isolate exogenous variation in the geographic distribution of ethnic groups. This empirical strategy relies on the assumptions that these historical trade networks are (i) predictive of current ethnic similarities between countries, and (ii) exogenous to current trade patterns. This chapter presents strong evidence in support of the relevance and exogeneity of the instrument: these pre-colonial trade routes track current ethnic distributions such as the Hausa along former West African networks or the Yao in South-East Africa, and were functionally disrupted by the process of colonisation which dismantled them in favour of different routes that were more favourable to the metropole. Threats to these assumptions include the potential for these historical routes to be tracking something other than the distribution of ethnic groups, such as colonial era infrastructure, resource endowments, or the persistence of preferences across time. Each of these is addressed through historical evidence or robustness checks.

Trade is affected by moral hazard. With imperfect contract enforcement and costly verification, buyers cannot easily ensure that sellers will correctly fulfil the agreed contract without incurring significant costs. This has been shown by Alfani and Gourdon (2012) and Greif (1993) which look at the cases of medieval European merchant collectives and Maghribi trade networks; with Guiso et al. (2009), Roy et al. (2014), and Yu et al. (2015) in more modern contexts.[2] Ties between ethnic groups can alleviate some of the costs associated with moral hazard as they allow the contracting agents to more effectively punish a partner who reneges on the deal through social connections. This has been developed along linguistic ties by Melitz (2008), Melitz and Toubal (2014), and Selmier II and Oh (2013)[3] with work on ethnic ties focusing generally on intra-country trade (Fafchamps, 2003; Robinson, 2016), with Aker et al. (2014) being an exception.

Related to this is the issue of information asymmetry which can lead to supply chain inefficiencies, signalling issues, and other trade frictions across borders. This is often addressed in the context of financial flows and optimal transactions (Portes et al., 2001; Portes and Rey, 2005), but it also impacts trade. Chae (2005) looks at the impact of information on timing and supply chains while Hudson and Jones (2003) and Rauch (2001) look at the issues of signalling and overcoming information barriers in cross-country trade.

More broadly under this umbrella is work relating to how social relationships and networks can influence economic activity. Fafchamps (2001) and Lorenzen (2001) both look at the role of social networks and clustering in reducing coordination costs for trade, while Burchardi and Hassan (2013) and Woolcock and Narayan (2000) look at the implications of these reduced costs for development. This is taken further still by Baldacchino (2005) which extends the role of social network coordination to the development of governance and economic institutions within communities. Also within this work on social relationships is literature concerning the role of imperial legacies and the social connections generated by shared history. Both Head et al. (2010) and Beestermoller and Rauch (2018) address the legacy and erosion of these imperial relationships between countries and their former metropoles. This is generalised by Gokmen et al. (2020) which looks at the long run trends related to the strengthening and weakening of imperial trade ties between countries.

Running parallel with this literature is work dealing with immigration, migrants, and economic activity which addresses similar themes, but from the opposite direction. These range from the more tangential literature on migrants and remittance flows (well summarised by Rapoport and Docquier, 2006) to the role of immigration in trade promotion. The role of different networks and the efficiency of certain networks compared to others are both ideas investigated by Felbermayr et al. (2010), Rauch and Trindade (2002), and Wagner et al. (2002). Differing effects on imports and exports is the focus of Head and Ries (1998) and Konecny (2011), while the historical impact of this effect is looked at by Dunlevy and Hutchinson (1999). This work frequently demonstrates that small groups of immigrants are sufficient to form strong trade links, an observation which echoes the finding of this chapter of the central role of minority ethnic groups in developing trade links between countries.

This chapter is divided into five sections. Section 2 introduces the ethnic variables,

[2] Also Spring and Grossmann (2016) for the opposite result.

[3] See Egger and Toubal (2016) for a summary.

while Section 3 presents the empirical strategy used in this analysis. The results are discussed in Section 4 and Section 5 concludes.

2 Data and Ethnicity

Trading relationships, particularly across borders, require information and trust to secure the transactions. Common ethnic ties may help in alleviating information asymmetries and reducing the threat of moral hazard. This chapter constructs two measures to investigate this relationship. The first captures the current ethnic links between countries based on population data. One issue with this measure is that current ethnic composition may reflect modern migration flows along trade routes. To mitigate this issue, I also construct a measure of common ethnic ties using historical trade routes.

2.1 Ethnic Variables

The data used to construct the ethnic variables come from a database maintained by the Joshua Project. The raw dataset is current for the year 2012 and, given the time-span of the analysis, significant population changes during the period are not expected across the sample. After combining and simplifying the raw data on ethnic groups, a total of 2381 distinct ethnic groups are included.[4] The data provide specific population counts of each ethnic group in a country, capturing the presence of population groups as small as 500 people.

Maps of historical ethnic territories are often suggested as the base to construct variables about ethnicity. Such maps are favoured because they remove the issues of endogeneity related to modern population movements (discussed further in Section 3.3). However, the information contained in such maps is related to the territorial extent of dominant groups within a region, this ignores the presence of smaller and minority groups which may be present across numerous regions, but never dominant in any. These smaller groups could potentially drive significant trade flows between countries dominated by different majority groups. In this respect, the use of detailed population data provided by this dataset is superior to more territory-based methods at identifying effects driven by potentially small groups.[5]

I construct four unique variables for ethnic links between countries for this analysis. The first is a straightforward dummy variable which takes the value of 1 if the countries share any ethnicity, and zero if they do not:

$$Link_{ij} = \begin{cases} 1 & \text{if } \exists\ k \text{ for which } \xi_{ik} > 0 \text{ and } \xi_{jk} > 0; \\ 0 & \text{otherwise.} \end{cases}$$

where ξ_{ik} is the percentage of the population in country i of ethnicity k. Across all country pairs the likelihood of sharing an ethnic group has a mean of 0.35. The threshold used for

[4]More detail about the data construction is provided in Appendix D.

[5]Variations on the main analysis using territorial maps are presented in Appendix A.

this indicator is any population regardless of how small, though the effective threshold is between 500 and 1000 people since that is the lower-bound on the population data. Within a search framework, this measure is analogous to a perfect search where if an individual of a given ethnicity exists in one country and there is a co-ethnic in the partner country, they will find each other with a probability of 1. These ethnic matches then generate connections that foster trade.

The second variable sums the number of ethnic links between countries. This incorporates the idea that the more shared ethnic groups a country-pair has, the more opportunities there are for ethnic networks to develop from one or more of those shared groups. This is as if each ethnic group were able to perfectly search for matches between the two countries, with the results of those matches being additive for trade promotion. This count of ethnic links is a simple sum of all shared ethnic groups in the countries:

$$Count_{ij} = \sum_k E_k$$

$$\text{with } E_k = \begin{cases} 1 & \text{if } \xi_{ik} > 0 \text{ and } \xi_{jk} > 0; \\ 0 & \text{otherwise.} \end{cases}$$

where ξ_{ik} is the percentage of the population in country i of ethnicity k.

A third variable captures the probability of an individual from one country randomly matching with a co-ethnic from the other country. This structure assumes implicitly that the meetings between traders are random (there is no directed search) and contrasts with the previous measure's perfect search structure. In this construction, the size of the relevant ethnicities matters to promote matches. The variable for this ethnic probability index is a sum of the population percentages across all ethnic groups in the two countries:

$$EPI_{ij} = \sum_k \xi_{ik}\xi_{jk}$$

where ξ_{ik} is the percentage of the population in country i of ethnicity k. Since the ξ terms are multiplied across the two countries, the value for a given ethnicity that is not shared between the countries is zero, since one or both of the ξ terms would be zero and a match would be impossible. These are then summed across all ethnicities and the resulting value is an index of the probability of making an ethnic match ranging from 0 to 1 with a higher value indicating greater probability of two co-ethnics meeting.

A final measure incorporates the extent to which the two countries trading are ethnically similar in composition. This implies the concept of co-ethnics in each country having similar tastes and habits that influence demand, fuelling trade in similar goods between countries. This ethnic commonality index is given as a sum of the differences of population percentages across all ethnic groups in the two countries:

$$ECI_{ij} = 1 - \frac{\sum_k |\xi_{ik} - \xi_{jk}|}{2}$$

where ξ_{ik} is again the percentage of the population in country i of ethnicity k. The ξ terms are subtracted across the two countries and then summed across all ethnicities in the country pair, resulting in an index of ethnicities shared by the countries ranging from 0 to 1 with a higher value indicating greater ethnic similarity.

2.2 Historical Networks as an Instrument

The instrument used in this analysis is based on pre-colonial trade routes constructed from textual data from a variety of sources on historical trade throughout the continent.[6] The collected data include both the location of trade posts and the connections between them. Ultimately, 196 separate trade posts are identified within five unique networks and one additional supra-network involving three of the smaller networks (see Figure 1.1). The five primary networks are: 1) North Africa; 2) the Congo Basin; 3) the East African Coast; 4) South-East Africa; and 5) the Horn of Africa. The larger supra-network of Central Africa includes trade posts within the Congo Basin, East African Coast, and South-East Africa networks where some connections across these networks existed.[7]

The trade posts in these networks are attested from the medieval period through to the nineteenth century and are quite consistent over time. The age of sources is somewhat skewed since most of the accounts are from Europeans and so networks that are further south are first mentioned later in the time range, coinciding with European arrival in the regions. Similarly, since the majority of included trade posts are observed and recorded by Europeans, this will naturally exclude many locally important but unobserved locations. This selection bias is difficult to overcome, but the impact on the instrument is of less concern because missed trade posts will bias the results against shared connections and so reduce the power of the estimation, meaning that the results may be underestimating the effect.

The routes that made up these trade networks emerged from the competition of various states and groups for resources and regional power. Trade along them consisted of a variety of activities, from formal long-distance caravans to individual merchants trading local products. The range of products traded along these routes were similarly varied, including raw materials, refined products, processed metals, manufactures of different types and qualities, cloths and textiles, art and jewellery, and agricultural goods.[8] The individual routes that made up the networks tended to be dominated by particular ethnic groups which traded along them and established communities in the various entrepôts.[9] This predominance of certain groups did not preclude other groups from sharing in the trade and different groups came to dominate different routes at different times.

The process of colonisation caused a breakdown in these routes and the influence of the groups that maintained them. Often there was a specific effort on the part of colo-

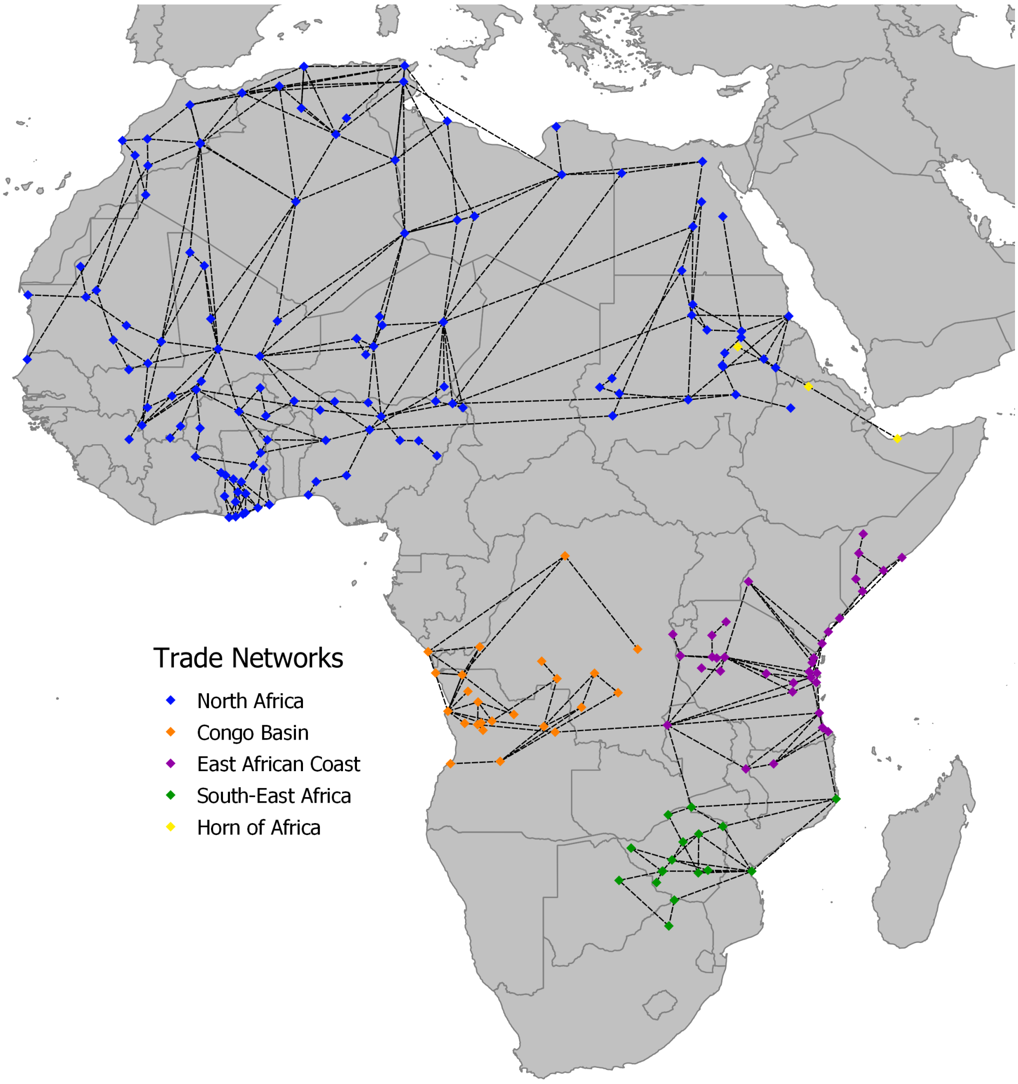
[6] Sources include Bovill (1995), Chirikure (2017), Coquery-Vidrovitch (2005), Gray and Birmingham (1970), King and van Zwanenberg (1975), and Ogot (1974).

[7] A more detailed description of the different networks and how they were constructed is presented in Appendix C.

[8] Further details on the networks and trade goods are provided in Appendix C.

[9] Some examples of this include the Yao in south-eastern Africa, the Nyamwezi near the Great Lakes, the Chokwe in the southern Congo Basin, the Hausa in the Sahel, and the Tuareg in areas of the Sahara.

Figure 1.1. Map of Historical African Trade Networks



Note: Major trade nodes correspond to the points and the dashed lines are the routes connecting them; current political borders are shown overlaid. (Source: Author's own calculations based on Bovill, 1995; Chirikure, 2017; Coquery-Vidrovitch, 2005; Gray and Birmingham, 1970; King and van Zwanenberg, 1975; Ogot, 1974).

nial powers to break the dominance of these groups and dismantle the trade networks in order to redirect resources to the colonial administrations.[10] When new infrastructure was created, it tended to serve a very different purpose than the previous trade routes; colonial-era infrastructure was either oriented towards extracting resources for the benefit of the metropole or directed by strategic and military concerns rather than commerce. This process followed an orientation that bypassed pre-existing trade posts and routes, displacing the networks and ensuring they remained that way. Colonial boundaries aggravated this process by severing the routes and restricting trade across borders—both between empires and sometimes within them. The subsequent fracturing of states along these restrictive borders during the process of decolonisation entrenched these divisions.

The result of this long process of colonisation and decolonisation was a significant disruption in the functionality of the pre-colonial trade routes by reducing their importance and often dismantling them altogether. Colonial infrastructure had different goals than these historical routes, and so modern transportation links mirror these strategic concerns rather than pre-colonial commercial ones. Despite this, previous movements of ethnic groups along these routes were not reversed, leaving in place the communities that had been established to maintain the trading system and allowing these networks to be an effective instrument for shared ethnic ties across countries.

The variable itself is a measure of the number of trade posts that a country pair possesses which share a network—effectively the number of shared network trade posts the country pair has within all networks:

$$H_{ij} = \sum_n (\delta_{in} + \delta_{jn}) R_n$$

$$\text{with } R_n = \begin{cases} 1 & \text{if } \delta_{in} > 0 \text{ and } \delta_{jn} > 0; \\ 0 & \text{otherwise.} \end{cases}$$

where δ_{in} is the number of trade posts in country i belonging to network n. This measures how integrated each of the two countries are within shared trade networks. In assigning network membership to trade posts, the path of least assumptions was followed—network membership was largely accepted at face value as attested in the sources with as few interpolations of wider linkages as possible. While this methodology increases the number of networks and the isolation of those networks, the benefit is that it leaves little open to interpretation beyond what is presented in the historical records. Empirically, by increasing the number of networks (and increasing the isolation of networks by under-measuring connectedness) it will bias the variable downwards, meaning that the results may be underestimated.

[10] The Portuguese campaigns against the Chokwe in the region of Angola and the British wars against the Asante in West Africa are examples of this process.

3 Empirical Strategy

3.1 Gravity Model

The objective of this chapter is to gauge the impact of common ethnic groups on the bilateral trade between countries. Gravity models have been used extensively in the literature to model the relationship between trade flows and transaction costs.[11] The structural gravity model, as presented in Anderson and van Wincoop (2003) and Feenstra (2004), relies on monopolistic competition and CES consumer preferences to represent trade between two countries as a function of the respective economic size of the countries and the obstacles between them. In its basic form, it can be written as:

$$X_{ijt} = K(Y_{it}Y_{jt})^{\alpha} \left(\frac{T_{ijt}}{P_{it}P_{jt}} \right)^{1-\sigma} \tag{1.1}$$

Where X_{ijt} represents the bilateral trade between country i and country j in period t, K is a constant, Y_{it} and Y_{jt} are the total outputs of country i and j respectively in period t, P_{it} and P_{jt} are the price levels of country i and j respectively in period t, and T_{ijt} represents transport costs between country i and country j in period t. Costs, T_{ijt}, are 'iceberg' transport costs such that $T_{iit} = 1$ and $T_{ijt} \geq 1$.

The trade cost term T_{ijt} is modelled as a composite (usually multiplicative) function of the various factors impacting the cost of trade,

$$T_{ijt} = f(\tau_{ijt}, d_{ijt}) \tag{1.2}$$

where d_{ijt} is the distance between countries i and j and τ_{ijt} is any other 'border effect' that could impact trade between i and j. This supplementary effect may often include transaction costs such as common legal systems, common institutional variables, contiguity, and any other dyadic impact on trade.

3.2 Empirical Specification

Using the multiplicative form of the T_{ijt} term, Equation 1.1 can be linearised with logs as follows

$$\begin{aligned} \ln(X_{ijt}) = \alpha \ln(Y_{it}Y_{jt}) + \ln(d_{ijt}) + \ln(\tau_{ijt}) \\ - ln(P_{it})^{\sigma-1} - ln(P_{jt})^{\sigma-1} + \epsilon_{ijt} \end{aligned} \tag{1.3}$$

Fundamental to the estimation of structural gravity equations is how to deal with the price indices P_{it} and P_{jt}, the unobservable 'multi-lateral resistance' terms (see Anderson and van Wincoop, 2003). I include country-year fixed effects to account for these and other unobservable factors.[12] Trade data are from the United Nations Conference on Trade and Development database, while the measure of distance used in this chapter is the linear

[11] See Head and Mayer (2014)

[12] In line with Glick and Rose (2002), Anderson and van Wincoop (2004), and Baier and Bergstrand (2007).

distance between the two capitals of the country pair.[13] The data covers the continent of Africa for the years 1995-2012, including North African countries and countries south of the Sahara without distinction.

In this analysis I divide τ_{ijt} into geographic variables, socio-historic variables, and economic variables. The geographic variables are sourced from the Centre d'Études Prospectives et d'Info Internationales (CEPII) database and include whether a country in the pair is landlocked ($land$), an island ($isle$), or whether the countries share a border ($bord$). The socio-historic variables include variables for common ethnicity ($ethn$), common official language ($lang$), and common colonial past (col). This last variable will capture any institutional similarities between the countries such as legal systems and governance structures. These variables are (almost always) time invariant but are not absorbed into the country fixed effects as they are country-pair based, allowing them to be separately identified from the country fixed effects. The final set of variables deals with broad economic ties and accounts for whether the countries are members of the same currency union ($curr$), customs union ($cust$), regional trade agreement (rta), or regional economic grouping (grp).[14]

With this division of the τ_{ijt} term and inclusion of the fixed effects, the specification of the model for estimation becomes,

$$\begin{aligned} \ln(X_{ijt}) = \alpha &+ \beta_1 \ln(d_{ij}) + \beta_2 land_{ij} + \beta_3 isle_{ij} \\ &+ \beta_4 bord_{ij} + \beta_5 ethn_{ij} + \beta_6 lang_{ij} + \beta_7 col_{ij} \\ &+ \beta_8 curr_{ijt} + \beta_9 cust_{ijt} + \beta_{10} rta_{ijt} + \beta_{11} grp_{ijt} \\ &+ \mu_{it} + \mu_{jt} + \epsilon_{ijt} \end{aligned} \tag{1.4}$$

where μ_{it} and μ_{jt} represent the country-year effects and ϵ_{ijt} is a normally distributed error term clustered at the country-pair (ij) level.

3.3 Instrumental Variable Framework

Using modern ethnic population data raises a significant concern about endogeneity such as reverse causation. It is possible that modern distributions of ethnic groups between countries are driven by modern trade flows, with people following the commerce to settle across borders in countries with strong trade links to their home countries. To address this concern, I use historical networks to instrument for shared ethnicity.

Since the instrumented variable (shared ethnicity) is a dummy variable and the primary instrument is not a dummy variable (shared network trade posts) I use the standard instrumental variable approach for continuous instruments with non-continuous endogenous

[13] For a more detailed discussion of distance measuring in gravity models see Mayer and Zignago (2011).

[14] Currency unions included in the data all use either a single currency or a pegged currency system in preparation for the introduction of a single currency. Customs unions include both common market agreements and more relaxed tariff agreement zones. Regional trade agreements are classified as agreements between countries that involve preferential trade zones or free trade zones. Economic groupings constitute agreements between groups of states that seek to further broad economic goals—most usually development or trade related objectives. Currency unions account for 3.4% of country pairings, 5% of country pairs are in customs unions, and between 26% and 29% of countries pairs belong to RTAs and regional economic groupings.

variables.[15] Following this method, I estimate a probit model first and generate predicted values of shared ethnicity. I then proceed with a normal two-stage least squares estimation with shared ethnicity instrumented by the predicted values from the probit model and the instrument (shared network trade posts).

The validity of historical networks as an instrument for common ethnic links rests on two assumptions. First, historical networks must be correlated with ethnic distributions. As discussed in Section 2.2, the unique nature of historical trade links along these routes meant that ethnic groups dominant along the routes would settle communities at various posts within the networks. This is strongly supported by the results of a naive first stage presented in Table 1.1. This table presents the results of regressing the various ethnic variables on the historical network instrument with country fixed effects included. From these results there is a strong association between geographic ethnic distribution and historical networks. Combining this with the historical evidence provides a strong case for these historical networks to be a predictor of ethnic distributions across countries.

Table 1.1. Naive First Stage

	(1) Link	(2) Count	(3) ECI	(4) EPI
Shared Network	0.023**	0.182**	0.031**	0.013**
	(0.003)	(0.031)	(0.006)	(0.004)
CFE	Yes	Yes	Yes	Yes
N	1396	1429	1429	1429
R^2	0.212	0.054	0.019	0.003

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the relevant ethnic variable; Shared Network is the number of shared network posts in the country-pair.

Added to this are possible measurement issues with the instrument. The potential of omitted routes or trade posts is ever-present in the data since it relies on historical textual accounts. While the reasons for this are more fully addressed in Section 2.2, the result is that there may be routes or trade posts that are missed in the data, either through limits to what those writing the accounts could see and understand or due to survivor-bias of the observed routes at the time of recording. Ultimately, this could reduce the number of recorded routes and trade posts in the data, therefore missing some connections between countries. While this problem is difficult to fix ex-ante, its effects are such that it would serve to under-report the extent of connections between countries and therefore under-represent shared ethnic groups, potentially reducing the predictive power of the instrument in the first stage.

The second assumption is that the exclusion restriction is satisfied: historical networks impact trade only through their effect on common ethnic groups. This assumption is

[15] As outlined in Windmeijer and Silva (1997).

threatened by the possibility that the instrument is actually capturing something other than the distribution of ethnic groups, such as infrastructure overlaid on the routes during the colonial period or historical endowments continuing to influence trade in products.

These possibilities are unlikely due to the nature of colonial governance which tended to construct infrastructure with the dual purpose of disrupting earlier economic systems and extracting resources from the country. This meant that colonial era infrastructure was often orthogonal (in purpose and construction) to earlier systems and networks for trade, concentrating resources for export to the metropole and limiting the flow of goods across imperial frontiers. Added to this is that the process of colonisation involved active efforts by colonising powers to disrupt and dismantle trade networks and connections through war and the imposition of strict borders.[16] Further, the process of decolonisation which solidified colonial era boundaries as national frontiers, erecting further barriers across what were by then defunct trade routes. This twin process of colonisation and decolonisation lead to the discontinuation of these pre-colonial routes, removing the potential for their direct influence on current trade flows between countries. A variable for common colonial past is included to further control for the possibility of some residual links existing between countries previously within the same empire, and robustness checks with a more subtle control for imperial legacies is presented in Section 4.3.

Table 1.2. Summary Statistics

Variable	Obs.	Mean	Standard Deviation	Minimum	Maximum
Link	16555	0.353	0.478	0	1
Count	16555	1.583	4.161	0	61
ECI	16555	0.010	0.035	0	0.428
EPI	16555	0.002	0.010	0	0.233
Network	16555	2.322	5.862	0	33

	Correlations				
	Link	Count	ECI	EPI	Network
ln(Dist)	-0.669	-0.584	-0.533	-0.342	-0.179
Border	0.431	0.696	0.542	0.330	0.248
Link	1.000	0.515	0.375	0.215	0.232
Count		1.000	0.585	0.288	0.226
ECI			1.000	0.792	0.105
EPI				1.000	0.025

Note: Link is the ethnic dummy variable; Count is the sum of shared ethnic groups; ECI is the Ethnic Commonality Index; EPI is the Ethnic Probability Index; Network is the number of shared network trade posts.

[16] A survey of the literature documenting this process is provided by Austen (1987).

Table 1.3. Comparative Statistics

	Ethnic Link=1	Ethnic Link=0	Difference
Ethnic Count	4.478	0	+4.478**
	(0.078)	(-)	
ECI	0.028	0	+0.028**
	(0.001)	(-)	
EPI	0.004	0	+0.004**
	(0.000)	(-)	
ln(Trade)	7.718	5.275	+2.443**
	(0.046)	(0.032)	
Language	0.618	0.533	+0.085**
	(0.006)	(0.005)	
Colony	0.459	0.313	+0.182**
	(0.007)	(0.004)	
	Network>0	Network=0	Difference
Ethnic Link	0.578	0.302	+0.276**
	(0.009)	(0.004)	
Ethnic Count	3.374	1.173	+2.201**
	(0.117)	(0.028)	
ECI	0.017	0.008	+0.009**
	(0.001)	(0.000)	
EPI	0.002	0.001	+0.001**
	(0.000)	(0.000)	
ln(Trade)	8.567	5.582	+2.985**
	(0.057)	(0.030)	
Language	0.643	0.545	+0.098**
	(0.009)	(0.004)	
Colony	0.477	0.339	+0.138**
	(0.009)	(0.004)	

* Significant at 0.10; ** Significant at 0.05

Note: This table shows the difference in means of Ethnic Link, Ethnic Count, ECI, EPI, ln(Trade), Language, and Colony when there is a common ethnicity between countries (Ethnic Link = 1) and not (Ethnic Link = 0); and when countries are part of the same network (Network > 0) and not (Network = 0); standard errors are in parentheses

3.4 Descriptive Statistics

Ethnic links between countries are not particularly common, the majority of country pairs do not have common ethnic communities; as the first panel of Table 1.2 shows, 35% of country pairs have a link. Added to this is the diversity of ethnicities across countries—no two countries are exactly the same in terms of ethnic composition as shown by the low maximum value of the Ethnic Commonality Index. Interestingly, no two countries have a common ethnicity that is a majority of the population in both countries; if this were the case, the maximum value for the Ethnic Probability Index would reach above 0.25.

The distribution of ethnicities across countries tends to be geographically clustered: countries that share an ethnicity tend to be closer together, as shown in the correlation panel of Table 1.2, and countries that share a border tend to share ethnicities, as shown in the same table. This is most pronounced for the Ethnic Count variable and Ethnic Commonality Index, the higher correlations of these variables with the border measure demonstrates that proximity is more important for country pairs to share multiple ethnicities and have similar ethnic compositions.

Despite the fact that the four ethnic variables are all correlated with each other, having a common ethnic link does not indicate that the countries are very similar ethnically. This is shown in the first panel of Table 1.3 where the values of both the ECI and EPI are low (0.028 and 0.004 respectively) when the countries share an ethnicity. This means that despite having a common ethnicity, the countries are neither similarly ethnically composed nor contain large groups of the same ethnicity. Importantly, this demonstrates that what is driving the common ethnicity between countries are minorities, not major ethnic groups or large populations. Even if an ethnic group makes up a large proportion of the population in one country this is not mirrored in any other country so regardless of the size of a given population in one country, the connections to other countries are still made through minority groups in those countries.

Added to this, country pairs that were part of the same historical network are more likely to share an ethnicity and tend to share more of them (3.4 groups on average) as shown in the second panel of Table 1.3. While these countries also tend to be more similar in ethnic composition, as with a shared ethnic link these values are not very large (1.7% for the ECI and 0.2% for the EPI). The ethnic link variable tracks connections between countries that are sometimes quite distant and the historical network variable appears to track a similar feature.

4 Results

4.1 Main Results

The first stage results of the IV analysis are presented in Table 1.4 for all four ethnic variables constructed in Section 2.1. The instrument performs well for only one of the variables considered, the Ethnic Link dummy variable, with the results for the other

Table 1.4. IV First Stage

	(1) Link	(2) Count	(3) ECI	(4) EPI
Shared Network	0.009**	0.041	-0.002	-0.000
	(0.002)	(0.026)	(0.006)	(0.004)
CYFE	Yes	Yes	Yes	Yes
N	16523	16523	16523	16523
R^2	0.684	0.547	0.469	0.221

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the relevant ethnic link variable; Shared Network is the number of shared trade network posts in the country-pair; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

specifications not being significant. In the context of the Ethnic Links variable, the Shared Network Variable is not found to be a weak instrument. The Ethnic Link variable captures the potential effect of very small populations within countries, rather than relationships conditional on the size of populations (as with the ECI and EPI measures) or the number of distinct communities (as with the Ethnic Count measure). The fact that the instrument performs well for the Ethnic Link variable but not the others demonstrates that the legacy of these historical trade networks is rooted in linking distant countries through minority groups. This makes the population-based data used in this analysis well suited to capturing this effect through the ability to capture small populations within countries.

Table 1.5 presents the results for the estimations with the dummy variable for a shared ethnic group. The first column contains the standard OLS specification. The coefficient is strongly positive and significant, with trade increasing by a factor of nearly two-thirds (64%). Importantly, the effect of common ethnic ties is significant with the inclusion of variables for common coloniser, common official language, or shared economic groups and treaties. Due to the geographic 'clustering' of what these variables capture it might be assumed that any effect of a common ethnicity would be capturing the regional nature of these variables. However, the results clearly demonstrate that the impact of the ethnic group variable is not as a proxy for other regional effects caused by the associations between governments.

The results from the instrumental variable estimation are presented in the second column of Table 1.5, with the reduced form in the third. The estimated IV parameter is equal to 0.82, indicating that trade more than doubles (an increase of 127%) between countries that share an ethnic link. This result is similar to the magnitude of customs unions and trade agreements, indicating that this is a powerful effect across countries. The large difference between the OLS and IV coefficients could be a result of two forces.

This could represent a local average treatment effect because of the possibility that

Table 1.5. Main Results

	(1) OLS	(2) IV	(3) Reduced Form
Ethnic Link	0.495**	0.820**	
	(0.163)	(0.313)	
Shared Network			0.036**
			(0.011)
CYFE	Yes	Yes	Yes
N	16523	16523	16523
R^2	0.451	0.690	0.451

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is a dummy variable; Shared Network is the number of shared network trade posts; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

the instrument is under-estimating the levels of connectedness between countries (due to the omission of some trade posts in the textual data). The instrument would then be measuring the effect of ethnic links on a particular sub-sample of the data with the most prominant historical networks where the effect of ethnic links could be stronger. This would produce larger results by conditioning the effect of a common ethnicity on the presence of a historical network. The other explanation is that there may be downward bias in the OLS specification (as indicated by some of the robustness specifications in Section 4.3). This would imply that the instrument is presenting a truer picture of the effect of ethnicity on trade than the OLS specification.

These results show that ethnic groups are playing a role in promoting trade between countries, likely by developing networks that can fill an institutional gap. There is a need for institutions to overcome barriers to trade like trust and information asymmetries, and these networks provide that. The power of these ethnic groups to drive trade is demonstrated in the strength of the effect, emphasising an impact beyond what may be expected from generally small communities. While the results indicate that ethnic groups are increasing trade across borders, the analysis is inconclusive about which ethnic groups are driving this effect and the extent to which the dominant historical ethnic groups remain so today.[17]

[17]While this issue is beyond the scope of this chapter, a preliminary analysis of the modern influence of several historically dominant groups is contained in Appendix A.

4.2 Alternative Measures of Ethnic Links

The results for the other ethnic variable constructions are presented in Table 1.6. The first column contains the results for the count of shared ethnic groups. The results are not significant, and the weak effect of this variable seems to indicate that the marginal impact of an increase in the number of ethnic links is low, if it were to impact it at all, with each additional link potentially correlating with a 2% increase in trade.

Table 1.6. Ethnic Variable Results

	(1)	(2)	(3)
Ethnic Count	0.023		
	(0.015)		
ECI		-0.096	
		(0.060)	
EPI			-0.033
			(0.049)
CYFE	Yes	Yes	Yes
N	16523	16523	16523
R^2	0.449	0.449	0.448

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

The second column presents the results for the ethnic commonality index which measures the degree to which the two countries are ethnically similar. The coefficient on the index is negative and not significant, indicating that a large population of a common ethnic group is not associated with an increase in trade. The coefficient, even if it were significant is quite small, with a 1% increase in the ethnic similarity of the two countries reducing trade by approximately 3%.

Additionally, when the ethnic link dummy variable is included, both are significant and the ethnic link dummy is positive with a similar magnitude to the baseline OLS. These results seem to indicate that while the presence of a common ethnic link fosters trade, as the population concentration of that link (and any other links) grows, the association is negative, implying a potentially non-linear functional form for the commonality index. Using a non-linear specification shows that trade is increasing in ethnic similarity until a critical threshold is reached and then begins to decrease. This supports the idea that small amounts of ethnic similarity (i.e. smaller ethnic groups that are common between the countries) are driving the effect.[18]

[18]The results of including the ethnic dummy variable and a deeper exploration of the non-linearity of the ethnic commonality index are provided in Appendix A.

Lastly, the third column contains the results for the ethnic probability index. The coefficient on this measure is negative, but it is not significant indicating that the random search concept of how ethnic groups influence trade is not appropriate in this instance. This result adds to the the result for the ethnic commonality index and lends support to the idea that the extensive margin (having a common ethnic group, regardless of size) is driving the effects on trade.

4.3 Robustness

I conduct a number of robustness checks on the baseline specification. I summarise the results in this subsection and include a more detailed discussion along with the relevant Tables and Figures in Appendix B. The robustness checks are divided into two parts: those dealing with the exclusion restriction for the instrument, and those addressing the measurement of variables and the sample.

Exclusion Restriction The first set of robustness checks address the exclusion restriction for the instrument discussed in Section 3.3. This includes the incorporation of variables to control for colonial institutions, colonial infrastructure, and resource endowments. Preferences are addressed through an analysis of results for specific product groups.

The instrument may be measuring the persistent effects of colonial institutions that currently impact trade through the harmonisation of administrative or institutional processes. While the baseline results include a variable for common coloniser, I introduce more nuance to the measure by interacting it with distance to account for the idea that colonial territories that were closer together tended to be under the same colonial administration or have more direct administrative ties. I then create a second measure of the common coloniser variable that controls for each different empire separately to address any difference in effect depending on the systems used in the different empires. These empire-specific variables are then interacted with distance to again capture the different effects depending on proximity. None of these measures affect the results, they all remain in line with those in Table 1.5, demonstrating that institutional connections formed during the colonial period are not influencing the instrument.

In order to rule out that the relationship is driven by common infrastructure built during colonisation, I create a variable using data from Jedwab and Moradi (2016) on rail lines constructed during the colonial period. I develop two separate variables for colonial infrastructure. The first is the presence of a colonial-era rail link between the countries. The second is the presence of a shared rail link interacted with having the same colonial power to incorporate the differing effects of rail lines that spanned imperial borders. The incorporation of these controls for infrastructure produce results with coefficients that are not statistically different from the baseline results. They remain positive and significant, indicating that infrastructure is not a factor in the relationship detected in the IV analysis.

Climate and topography greatly influence what agriculture or raw materials can be produced within a country and these resource endowments can influence both historical and modern trade. To show that the baseline results are not related to the distribution of resources, and more directly that the instrument is not influenced by these endowments, I create a series of variables capturing the different climatic and topographical distinctions between countries. These include things such as soil quality, rainfall, and elevation. Each variable is constructed by taking the difference between the measured value of the relevant aspect for each country, producing a continuous index of difference between the countries—effectively measuring the climatic and topographical 'distance' between countries. By capturing these differences, the variables can approximate differences in resource endowments which could drive trade between countries. The inclusion of these controls do not produce results that are significantly different from the main results and show that the instrument is robust to differences in resources and endowments.

To demonstrate that the instrument is not influenced by the effect of preferences established through pre-colonial trade and still currently influencing trade flows, I estimate the relevant product categories separately. If the effect were driven by the persistence of historical consumer preferences, there would be a differentiated effect between product categories that include goods influenced by these preferences—products traded historically such as agricultural produce and raw materials—and product categories developed after these preferences were established (such as manufactures, chemical products, and machinery). The results do not show such a differentiation, making it very implausible that preferences are driving this effect. There is very little difference between product categories: all the results are in line with the baseline estimation for each category with the exception of fuel which is a positive outlier.

Ultimately, the resilience of the results in the IV analysis to the inclusion of variables for institutions, infrastructure, or resource endowments, combined with the robustness of the results across product categories provide support for the exclusion restriction assumption and that the impact of historical trade networks on modern trade flows is through the effect of ethnic groups.

Measurement and Sample The second set of robustness checks look at variations in the measurement of variables and changes to the estimation sample.

The first of these checks is on the construction of the instrument used in the IV analysis. To test the sensitivity of the results to the construction of the instrument, I create an alternative measure of shared network posts incorporating the size of the various trade posts. The value for this measure is based on the number of times each trade post is mentioned in the sources, approximating the importance of the various trade posts. I then use this to weight the sum of trade posts across networks, giving more prominence to countries with trade posts that are more important within the networks. I use this new weighted version of shared network posts to instrument for shared ethnic links. It performs in line with the original construction and produces no appreciable change in the results, indicating the strength of the historical data as an instrument.

As a second test I include measures that account for religious ties between countries. Since often religion and ethnicity have a strong interdependence and both can influence networks that promote trade, incorporating religion controls helps to confirm the unique effect of shared ethnic groups beyond (and possibly across) religious lines. Three variables are constructed in line with those created for ethnicity in Section 2.1: a dummy variable for shared majority religion, a proportional index, and a measure of religious similarity between the countries. The inclusion of these measures does not change the results, demonstrating that the effect of ethnicity is robust to religious effects. To emphasise this, an identical exercise is performed for Muslim populations only due to the historical prominence of Islam within trade relationships on the continent. The effect of ethnic links is also robust to the inclusion of these variables.

Third, because of the subjectivity of ethnic categorisation, I test the robustness of the results to different constructions of the main ethnicity variable. I increase the population threshold to 1000 people and 10000 people, and also change the definition of an ethnic group, lumping some groups together to reduce the overall number of groups. The results are not significantly impacted by these changes, indicating that the results are not sensitive to the threshold imposed for a common ethnicity or the definition of specific ethnic groups.

As a fourth check, I explore variations in the measure of distance in the gravity equation used in the estimation. I include three alternative measures of distance: the distance between major cities, the distance between country centroids, and a distance measure based on transportation methods that measures the sea and road distance necessary to connect two countries. The first of these substitutes the largest city in the country for the capital which is used in the original analysis. The second uses country centroids instead of cities. The third measure is a construction that measures the total distance along roads and waterways between the countries, giving a better approximation of the actual distance the goods travel. The results remain unchanged with the distance as measured between major cities. The incorporation of the other two measures produces coefficients for the ethnic links variable that are slightly lower in magnitude, though they are not statistically different from the main results.

Fifth, I accommodate the inclusion of zero trade flows and missing data through the use of the Poisson pseudo-maximum likelihood (PPML) model of Silva and Tenreyro (2006). Due to the low number of zeros in the data (0.009% of cases) and the comparatively larger number of missing data points (31.6% of cases) two different constructions are used. The first includes only those zero flows identified as such in the data. The second converts the missing data to zeros and includes those observations as well since often missing values can be misinterpreted zeros. Both these specifications produce results in line with the baseline estimation in column 1 of Table 1.5. In both cases the coefficients are slightly larger indicating that the OLS specification appears to be under-estimating the effect.

Finally, I test the sensitivity of the results to sample choice. The first round of these removes various regions from the sample that may have a disproportionate affect on the results either due to having a small number of ethnic groups distributed across many countries or a high number of ethnic groups shared among many smaller countries. Removing these regions changes the magnitudes of the coefficients, but all remain strongly positive and significant. Despite the fluctuations in the magnitudes, the coefficients are not statistically different from the baseline results. The second round performs tests on the time period included in the sample by estimating year cross-sections. The coefficients for the individual years vary somewhat, but all except one are in line with the main result. The cross-section for 2001 produces an anomalous negative result that could be explained by a significant drop in underlying global trade trends during that year. Results for cross-sections using multi-year averages are broadly similar, with coefficients that are positive and not statistically different from the baseline estimates. These exercises demonstrate that the results are not overly sensitive to sample choice.

5 Conclusion

This chapter provides evidence for the effect of ethnic ties on trade between countries. I build a dataset based on modern population distributions and construct several measures to capture various types of ethnic link. The research design enables the use of exogenous historical trade networks to instrument for modern ethnic group distribution across African countries in order to provide verification of causal mechanisms. I find that ethnicity has a large and significant effect on trade between countries, countries that share an ethnicity more than double their bilateral trade. This result is robust to different specifications of the ethnic variable and variations in the sample and controls. Ultimately, the results indicate that the defining feature of the effect is on the extensive margin—having an ethnic link is more important than the size of populations or the number of shared links. These results imply that the presence of common ethnicities goes some way towards mitigating issues related to moral hazard and asymmetric information, facilitating trade across borders.

This study does not come without limitations. Ethnic issues are consistently a point of debate with respect to Africa and the ethnicity variable identified in this analysis could always be adjusted to capture different degrees of variation between ethnic groups. The method chosen here is effective and seeks to maintain localised ethnic self-identification. Similarly, the issue of a difference in effect between discontiguous and contiguous ethnic groups is beyond the scope of this study but would offer interesting avenues to explore the ways in which ethnicity builds trade bridges between countries. The instrument used in this analysis comes with its own qualifications (discussed more fully in Section 3.3) and improving on its construction would allow for more insight into the distribution of ethnic groups and the potential persistence of the impacts of specific historical groups on trade.

Despite these issues, the insights provided here are extremely relevant to research in trade and the impacts of local networks on trust and information. Beyond this, the unique instrument used opens avenues into economic history with the legacy of historical trade and the distribution of dominant ethnic groups. This work also provides evidence at a broad scale of the ethnic border effects found in several studies of cross-border market prices. This contributes to the growing consensus around the importance of local networks for trade and provides generalisable results for these effects.

Although not necessarily obvious, the generalisability of the results lend themselves to possible policy implications. While formally exploiting ethnic networks is difficult in a policy context, these results reveal potential gaps in contract and information quality that are easily solved with policy measures. Pursuing reforms in these areas would reduce the need for ethnic networks to mitigate trust and information issues and would encourage the development of more robust, institutionalised trade between countries.

A Appendix: Alternative Specifications

A.1 Non-linearity of Ethnic Commonality Index

As discussed briefly in this chapter, the effect of the Ethnic Commonality Index is negative and insignificant. There is evidence, however, that the effect of this variable is non-linear and that it can have a significant impact on trade between countries and supports the conclusion of the principal analysis regarding small ethnic groups. This subsection presents these non-linear results for the ECI variable.

Table 1.7 shows a more in-depth presentation of the results for the ECI variable. In the first column is the basic regression for reference. The main coefficient of the ECI variable is negative and insignificant. This changes when the dummy ethnic link variable is included in the specification. As shown in column 2, the coefficient for ECI remains negative but becomes significant while the coefficient for the Ethnic Link variable is consistent with the results presented in this chapter.

Table 1.7. Non-linear ECI

	(1) ECI	(2) With Dummy	(3) Non-linear
Ethnic Link		0.498**	
		(0.163)	
ECI	-0.096	-0.098*	-0.352**
	(0.060)	(0.059)	(0.105)
sqrt(ECI)			0.374*
			(0.144)
CYFE	Yes	Yes	Yes
N	16523	16523	16523
R^2	0.449	0.452	0.451

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is a dummy variable; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

This would seem to indicate that there may be a non-linear effect in terms of the ethnic commonality variable; with the presence of a link creating a positive effect on trade, while the impact decreases as ethnic similarity increases. In order to explore this possibility a non-linear model is developed for the ECI variable of the following form:

$$\ln(X_{ijt}) = \alpha + \beta_1 ECI + \beta_2 \sqrt{ECI} + \beta_3 D_{ij} + \beta_4 M_{ijt} + \mu_{it} + \mu_{jt} + \epsilon_{ijt}$$

where D_{ij} and M_{ijt} are time invariant and variant (respectively) control variables used in the principal analysis. Country-year fixed effects are represented by the μ terms and ϵ_{ijt}

is an error term, clustered at the country pair level.

The non-linear specification (presented in column 3) has the ECI variable significant in both terms of the polynomial. The negative coefficient on the ECI term and the positive coefficient on the sqrt(ECI) term gives the function an inverted-U shape. This means that the effect of the ECI variable on trade is increasing for low levels of ethnic commonality between the countries until a critical value is reached and the effect decreases again.

This result reinforces the notion that the effect of ethnic groups on trade is one of minority groups, with small shared groups driving most of the effect. The ECI result adds to this by demonstrating that the effect is more prominent in largely heterogeneous countries (with a low overall ECI score) that share some small groups (increasing with the ECI variable for lower values).

A.2 Ethnic Maps

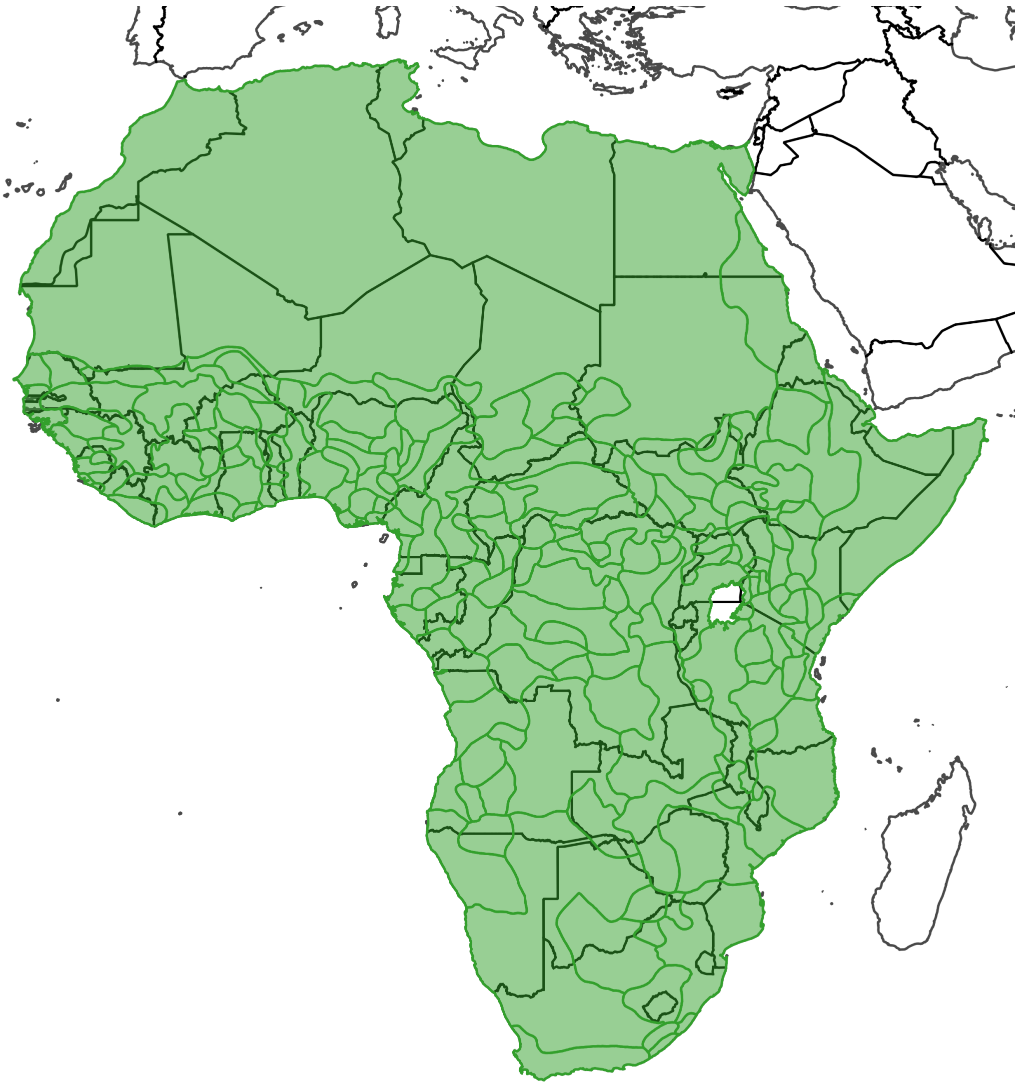
While the main analysis of this chapter uses detailed population data in order to capture the effect of small minority groups within countries, there are alternative ways to measure the presence of common ethnicity. One dominant method is the use of maps of historical ethnic territory. While the use of these maps removes the ability to measure the effect of small ethnic groups, they also remove the issue of endogeneity by determining common ethnicity without the potential influence of modern trade flows. Two maps are presented in this appendix to construct similar measures of shared ethnic links.

The first map (see Figure 1.2) is from the Times of London and was published in 1972 (Lewis and Foy, 1972) while the second map (Figure 1.3) is from Murdock's 1959 study of ethnicity in Africa (Murdock, 1959). The use of these representations of common ethnic communities is straightforward, as these maps represent clear ethnic distributions. However, it also creates two issues: 1) the structure of these maps and the territories they illustrate remove the ability to assess the impact of the size of populations within countries as there is no way to determine how densely populated the various parts of the mapped territories were; and, 2) it clouds the potential influence of smaller minority groups that are not represented as dominant territorially. This creates an increased sensitivity to classification issues since only groups that were considered the dominant players in various territories have been included in the maps and it necessarily reduces the number of ethnic groups (to the hundreds rather than thousands as in the population data). In this sense, they can only be used as checks on the main analysis.

In the case of each map, the map and its ethnic territories were digitised[19] and overlaid on a map of modern African countries. Any pair of countries that share the same ethnic territory (or share parts of a discontiguous ethnic territory) were deemed to have a common ethnicity. This dummy variable operates in the same fashion as that in the principal analysis and is independently created for each map. Both maps suffer from issues of the clumping and splitting of groups in different ways. While this problem is somewhat

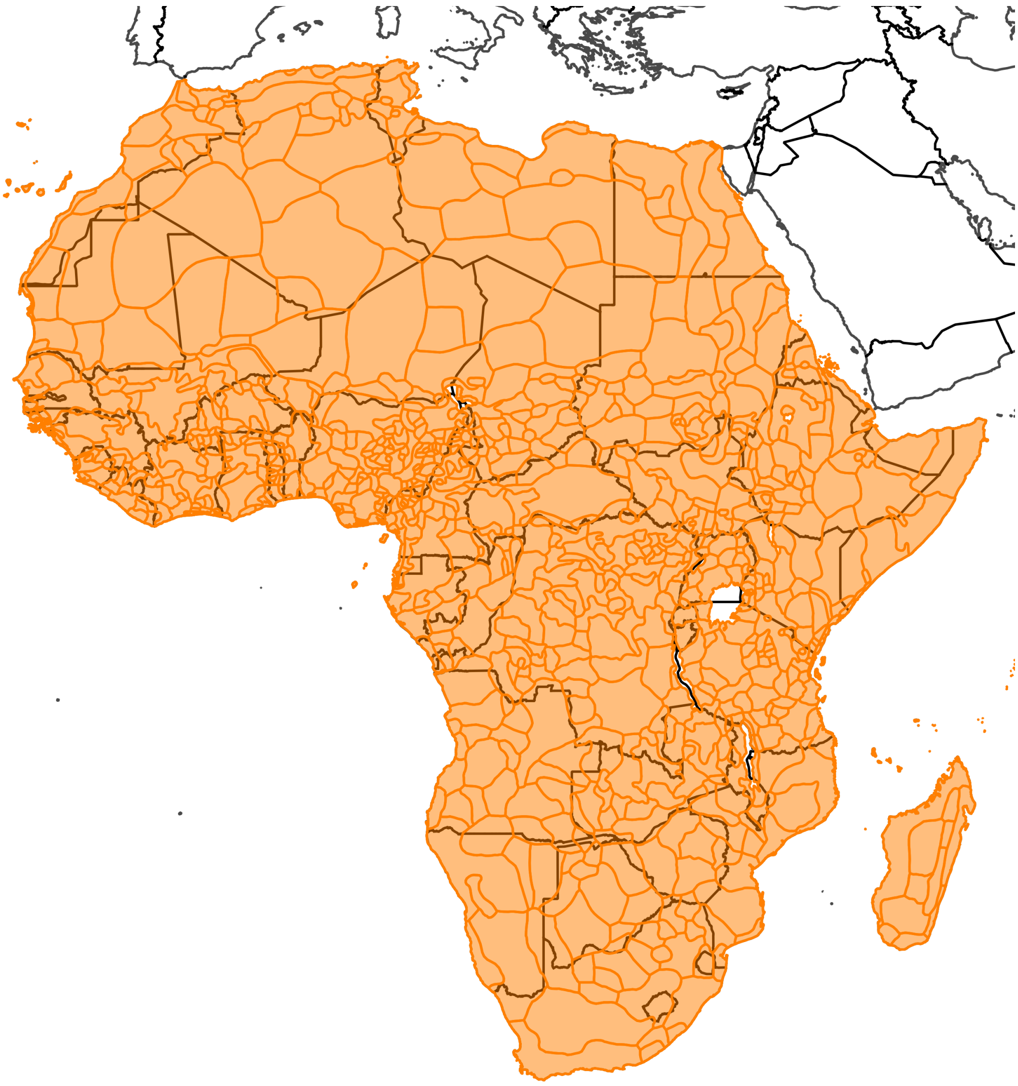
[19] Digitisation of the Murdock map comes from Nunn (2008), the Times map was digitised by the author.

Figure 1.2. Digitised Map of African Ethnic Groups (Times of London, 1972)



Note: Colour blocks correspond to areas occupied by individual ethnic groups as defined in Lewis and Foy (1972); current political borders are shown overlaid.

Figure 1.3. Digitised Map of African Ethnic Groups (Murdock, 1959)



Note: Colour blocks correspond to areas occupied by individual ethnic groups as defined in Murdock (1959); current political borders are shown overlaid.

mitigated by the fact the variable for common ethnicity is constructed as a dummy for the existence of any common ethnicity, it makes it difficult to conduct a more detailed analysis. Splitting is more problematic as the structure of the dummy variable for common ethnicity cannot mitigate this impact; in fact, the problem is likely exacerbated by the nature of the common ethnicity variable. For this reason, these must be viewed largely as robustness checks on the main analysis. The Times map produces a total of 167 unique ethnicities, with 17.8% of country pairs sharing an ethnic group, while the Murdock map has a total of 835 unique ethnicities, with 12.3% of country pairs sharing an ethnic group.

Times of London Map, 1972 The first map used to derive ethnic data is the map of ethnic territories published in the Times of London in 1972 (see Figure 1.2). This map was collated from several disparate sources and it is quite comprehensive in its treatment of African ethnic group territory. There is only one significant issue with its representation: the map tends to clump together some groups that the authors decided (somewhat arbitrarily it appears) belong together. This is most clear among the nomadic and semi-nomadic peoples of the Sahara which have been clustered into one pan-Saharan mega-group that also includes almost all of North Africa. While the motivations for this are a mystery, it does unfortunately mean that all Saharan and North African ethnicities are treated as a single group (lumping together peoples as disparate as Arabs, Tuareg, and Sahrawi). This issue is also evident in other places, though less clearly so. In several parts of the map some groups are included together while other groups are excluded (for example, Pedi is clustered with Eastern Sotho but Northern Sotho is excluded).

Murdock Map, 1959 The second map used for ethnic data is the map included with Murdock's 1959 study of ethnicity in Africa (see Figure 1.3). The entire work is a study of African ethnicity and culture that included a comprehensive map of the locations of various African ethnic groups. Murdock's map is very comprehensive, encompassing the territorial limits of hundreds of ethnic groups. In contrast to the Times Map, this map tends to suffer from the error of splitting (Braun and Hammonds, 2008). Ethnic groups that would normally be considered one group are divided into separate parts in seemingly arbitrary ways. The most notable example of this is the treatment of the Somali clans. In Murdock's map, each clan is classed as a separate ethnic group, despite them all being ethnically Somali. This would be equivalent to deciding that the Duncans and MacDonalds of Scotland are separate ethnicities. The issue with this is not the number of ethnic groups, but rather the way in which they were split. The splits that were made seem to manufacture new divisions within ethnic groups that the members of the groups may not recognise themselves.

Results The first column of Table 1.8 presents the results of the estimation using a common ethnicity variable constructed from the Times of London map (see Figure 1.2). The common ethnicity variable is strongly positive with a coefficient of 0.886, representing a 130% increase in trade. This is in line with the results from the IV estimation in the main analysis. The second column of Table 1.8 shows the result of using the Murdock map (see Figure 1.3) to construct a variable for common ethnicity. Again, the sign and the magnitude of the ethnic variable is in line with the previous estimation.

Table 1.8. Map-derived Ethnicity

	(1) Times	(2) Murdock
Ethnic Link	0.886**	0.860**
	(0.197)	(0.284)
CYFE	Yes	Yes
N	16523	16523
R^2	0.693	0.691

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is a dummy variable; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

These results provide further support for the idea of local ethnic networks facilitating trade between countries while simultaneously tackling the issue of population movements. By using maps of the historical extent of ethnic groups on the continent, the issue of modern trade driving population movements does not impact this estimation. The maps themselves approach the idea of ethnicity very differently, with the Times map tending to clump ethnicities together, while the Murdock map tends to split them. The strength of the results in the face of these divergent approaches to classification lends further support to the case for common ethnicity driving trade between countries.

A.3 Specific Groups

While the historical persistence of the influence of specific ethnic groups is beyond the scope of this chapter, this subsection includes some preliminary investigation of this question by looking at the modern influence of four historically important groups. Three of these groups (the Fulbe, Hausa, and Tuareg) are from the Sahel-Sahara area while the fourth, the Swahili, are present in East Africa. Below are brief historical sketches of each group's impact on trade and commerce in its region followed by the results.[20]

Fulbe The Fulbe people (alternatively called Fulani, Fula, Peul, among many others depending on the country and language) are a semi-pastoral ethnic group from West

[20] For a more complete discussion of historical linkages between cities in the Sahel, Sahara, and East African coast see Bovill (1995), and Coquery-Vidrovitch (2005).

Africa. They are one of the most widely dispersed ethnic groups in Africa and are the largest group in the Sahel region. Their origins are from the northern Senegambian region but through migrations, conquest, and trade they spread throughout West Africa and the Sahel. Historically, Fulbe states in the Sahel began to assert themselves in the sixteenth century as they laid claim to trade routes, entrepôts, and merchant towns. As these states became more successful and prosperous, Fulbe from surrounding regions increasingly abandoned historically pastoralist practices and settled in urban centres, often converting to Islam. Fulbe merchants frequently travelled the caravan routes in the region, setting up communities in various cities throughout the area. Fulbe states grew in dominance in the region, eventually culminating in the Sokoto caliphate of the nineteenth century.

Hausa The Hausa people are concentrated in the eastern Sahel region and have a long history of independent city states commanding trade links across the Sahara. The Hausa city-states established themselves as southern terminals to trans-Saharan trade and Hausa merchants plied these trade routes frequently. By the twelfth century the principal Hausa city states of Katsina, Zazzau, Gobir, Kano, Rano, and Biram had come to dominate Sahelian and trans-Saharan trade. The mercantile power of the cities extended far along rivers and into the desert, though the territory they directly controlled remained smaller.

Tuareg The Tuareg people principally inhabit the Sahara and are traditionally pastoralists. The nomadic lifestyle of many Tuareg people in the Sahara led them to be relied upon as caravan drivers through the desert. Since antiquity these caravans had been organised and led by Tuareg people and the Tuareg came to dominate the trade routes and way-stations within the desert.

Swahili The Swahili people are from the eastern coast of Africa, the word Swahili is derived from the Arabic term meaning ‘people of the coast’. The Swahili Coast has been a principal trade hub since antiquity, it was called Azania in the Greco-Roman period and had been attested to in numerous sources throughout history. The eventual rise of Swahili city-states on the coast and settlements on offshore islands was a direct result of the prosperity created by the trade networks of the Indian Ocean. These cities would serve as points of exchange, buying goods from the interior and then selling them onwards to their networks in the Indian Ocean. Swahili communities emerged at both ends of this trade network; establishing communities in the interior of East Africa and at the terminals for their Indian Ocean trade. As a result of this trade the cities of Bagamoyo, Kilwa, Mzizima (now Dar es Salaam), Malindi, Mombasa, Pangani, Quelimane, Sofala, and Tanga became some of the richest cities in the region. These networks were so lucrative that they attracted first the Portuguese and then the Omanis, who dominated the region from the sixteenth century.

Results Table 1.9 presents the results for the four relevant models: column 1 presents the impact of Fulbe networks; column 2 shows the effect of Hausa networks; column 3 deals with Tuareg networks; column 4 includes the impact of Swahili networks. Each specification includes the normal Ethnic Link variable from the main analysis and a variable for the specific group (Fulbe, Hausa, Tuareg, and Swahili). This helps to isolate the effect of the individual group under consideration from the more general average ethnic link effect across the rest of the continent.

In the first three columns, the variable for the specific group is positive, but not significant, indicating that the effect of the given ethnic groups cannot be separated from the general effect of having a common ethnicity. The effect for the Swahili group is strongly positive and significant, indicating that the effect of Swahili networks is separate from and stronger than the general ethnic link effect. These results present some preliminary evidence for the potential persistence of historically dominant groups, but they are not comprehensive. Further analysis of more groups, ideally with more localised impacts, would shed more light on this issue.

Table 1.9. Historical Ethnic Groups

	(1) Fulbe	(2) Hausa	(3) Tuareg	(4) Swahili
Ethnic Link	0.459**	0.473**	0.492**	0.332*
	(0.168)	(0.173)	(0.163)	(0.176)
Specific Group	0.232	0.083	0.200	0.690**
	(0.242)	(0.255)	(0.333)	(0.292)
CYFE	Yes	Yes	Yes	Yes
N	16523	16523	16523	16523
R^2	0.452	0.451	0.451	0.453

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link and Specific Group are dummy variables; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

B Appendix: Sensitivity Analysis

This appendix presents a series of robustness checks around the baseline specification.

B.1 Exclusion Restriction

The IV analysis presented in this chapter rests on the assumption outlined in Section 3.3 that historical trade networks only impact modern trade flows through their influence on ethnic links between countries. While this assumption is supported by historical accounts, there is still potential that other factors such as infrastructure and institutions created during the colonial period, country resource endowments, or consumer preferences could be related to these former trade routes and the legacy of these factors is what is actually being picked up in the analysis. To address this, I create a series of variables to control for colonial infrastructure, institutions, and country endowments. To address the concern around consumer preferences, I analyse the effect of the instrument on different product categories.

Institutions and Infrastructure To control for institutions and infrastructure created during the colonial period, I develop more nuanced variables for common coloniser effects. To account for institutions, I construct three additional variables for the effects of a common coloniser. The first of these is an interaction term of the colonial variable from the baseline estimation and distance to proxy for potential colonial institutional effects. Former dependencies of the same coloniser that are closer together could be more likely to have been under the same administrative umbrella or to have had shared institutional connections with each other. This potential is captured by this variable. The effect of this control in presented in column 1 of Table 1.10. The coefficient for common ethnic groups across countries remains unchanged with this inclusion.

The second of these institutional variables is a more refined measure of the common coloniser variable which splits the variable into control variables for each of the six major colonial powers: Belgium, France, Italy, Portugal, Spain, and the United Kingdom. This division allows for different institutional effects depending on the specific imperial power. The inclusion of these controls in the second column also do not affect the results of the variable for a common ethnicity. The third variable takes the separated colonial variables and interacts them with distance as in the first case. This modification does not significantly impact the effect of the Ethnic Link variable.

To tackle the effect of infrastructure, I create a variable to account for the presence of colonial railroads in order to incorporate controls for infrastructure effects, shown in the fourth column of the same table. The data for this dummy variable is from Jedwab and Moradi (2016) and it accounts for the presence of the same rail system within both countries. Railroads were the primary form of connective infrastructure created during the colonial period and this variable gives an indication of the presence of infrastructure links between countries. The presence of a common rail system does not negatively impact the results for a common ethnic link, lending support to the assumption that the instrument is indeed independent of any colonial infrastructure effects.

Table 1.10. Colonial Institution Controls

	(1) Colony-Dist	(2) Coloniser	(3) Coloniser-Dist	(4) Rail	(5) Colony-Rail
Ethnic Link	0.818**	0.879**	0.835**	0.863**	0.791**
	(0.304)	(0.309)	(0.300)	(0.318)	(0.332)
CYFE	Yes	Yes	Yes	Yes	Yes
N	16523	16466	16466	16523	16307
R^2	0.691	0.693	0.695	0.691	0.679

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network; Rail is a variable controlling for the presence of a shared colonial railroad (data from Jedwab and Moradi, 2016); ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

In the final column of Table 1.10, I create an additional measure of colonial infrastructure by interacting the variable capturing the presence of a rail system with the common coloniser variable. This captures the dual presence of a rail system that is under the control of the same colonial power in order to differentiate from rail systems that spanned empires. The inclusion of this more subtle measure of colonial infrastructure does not significantly change the overall effect of having an ethnic link, it is not statistically different from the baseline estimation and remains strongly positive and significant. These results give empirical credibility to the historical evidence supporting the independence of the instrument.

Resource Endowments In order to control for the potential impacts of country endowment, I create a series of variables to account for different climatic and topographic features of the various countries in the dataset. While endowments of resources are difficult to calculate for historical periods (particularly pre-colonial Africa), climate and topography differences can proxy for such endowment effects between countries: if two countries have significantly different climates or topographies they are likely to possess significantly different resources which could drive historical trade patterns that are mirrored in modern trade flows due to the persistence of these resource differences.

I create five separate variables to control for different geographic features: soil quality, forest cover, rainfall, elevation, and mountains. Each variable is calculated as the arithmetic difference between the values of the measures for the two countries in order to capture the climatic or topographical 'distance' between the environments of the two countries.[21] The results are presented in Table 1.11. The data for all but one of these variables comes from the World Bank, with the measure used to create the mountains variable derived from the ruggedness variable used in Nunn and Puga (2012).

Table 1.11. Country Endowment Controls

	(1) Soil Quality	(2) Forest Cover	(3) Rain	(4) Elevation	(5) Mountains	(6) Combined
Ethnic Link	0.859**	0.856**	0.872**	0.902**	0.827**	0.894**
	(0.327)	(0.328)	(0.320)	(0.322)	(0.324)	(0.324)
CYFE	Yes	Yes	Yes	Yes	Yes	Yes
N	15800	15800	16523	16523	16523	15800
R^2	0.696	0.696	0.691	0.690	0.691	0.697

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network; soil quality, forest cover, rain, come from World Bank data; mountains variable comes from Nunn and Puga (2012); ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

The first column presents the soil quality variable. This control measures the difference in the percent of arable land in the countries, capturing the variation in soil quality and potential agricultural output for the two countries. The inclusion of this control does not change the results for the Ethnic Link variable.

The second column shows the results for the forest cover variable which measures the difference in the percent of forest coverage in the two countries. This captures the difference in availability of resources such as timber and plant material, but also proxies for the prevalence of open land such as savannahs and deserts. Including this variable does not appreciably change the coefficient on Ethnic Link or its significance.

The third column includes a variable measuring the difference in annual rainfall between the countries. This control captures the differing potential for crops and resources based on water requirements, since areas that are more wet are able to produce different crops than drier regions. There is no significant change to the results with the inclusion of this measure.

The fourth column adds a variable which captures the difference in average elevation of the two countries. This accounts for the fact that different elevations produce different micro-climates that influence available resources and agriculture. The is no significant difference in the results when this variable is included, despite the small increase in the coefficient on the Ethnic Link variable.

The fifth column includes a measure of the difference in ruggedness between the two countries. As mentioned, this value is adapted from Nunn and Puga (2012) and quantifies the heterogeneity in the topography of a country—essentially the extent and frequency

[21] Using non-linear transformations of the variables that weight differences depending on their place within the variable distributions does not produce different results (results available upon request).

of changes in elevation. The variable used here captures the difference in ruggedness between the countries since different resources tend to be present in more mountainous or hilly environments. The inclusion of this measure also does not affect the results, with the coefficient on the Ethnic Link variable almost exactly in line with that presented in Table 1.5.

Finally, in the last column, all five measures for different climatic and topographical effects are included together as collective controls for endowment differences. The results for the Ethnic Link variable are robust to the inclusion of these different controls, demonstrating that the instrument is not picking up persistent differences in resource endowment.

Preferences It is possible that the instrument is actually capturing persistence in consumer preferences across time, where habits and tastes established prior to the colonial period create a tendency for consumers to buy specific goods from specific countries. This could be reflected in historical trade routes and the persistence of these preferences could be influencing modern trade flows. While preferences themselves cannot be measured directly for the pre-colonial period, tastes can be assumed from the types of goods traded. In theory, if these potentially static preferences are influencing modern trade flows then there should be some consistency in the types of goods traded. If this is the case, then agricultural goods and raw materials should be driving the effect of common ethnicity and this preferences effect is what the instrument will be estimating.

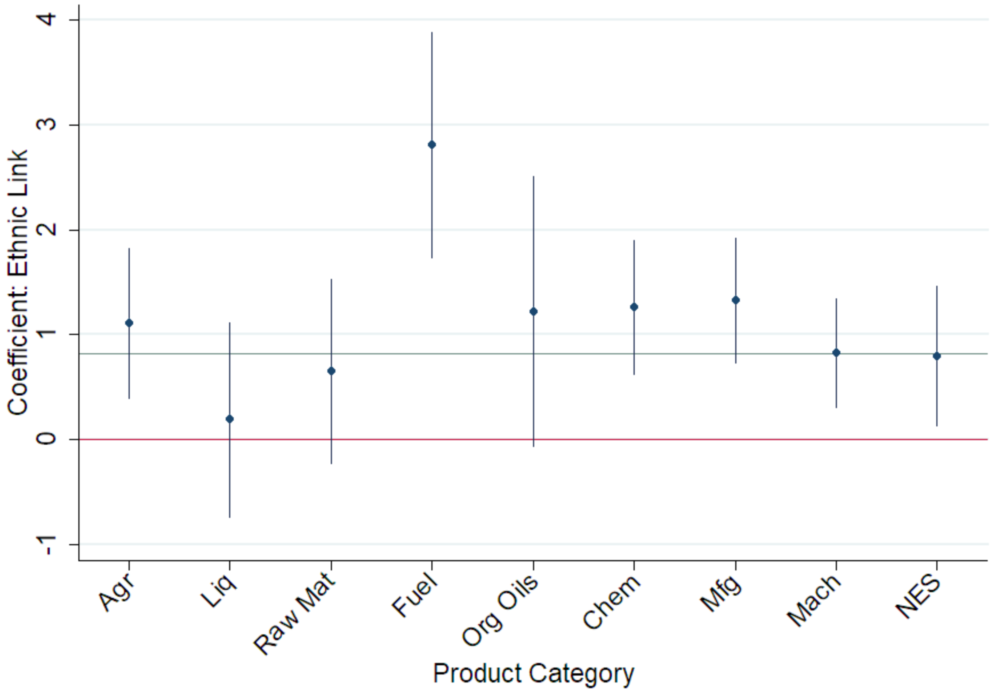
Figure 1.4 presents the results for estimations using specific product categories. The categories used (in the order they appear on the graph) are: agricultural products; liquor, beverages, and tobacco; raw materials; fuel; organic oils and fats; chemical products; manufactures; machinery and vehicles; and products not elsewhere specified. The product categories used are standard SITC one-digit categories.[22]

From the graph it is clear that all product categories are in line with the results from the main analysis (blue line) except for fuel which is an outlier. This shows that the effect of a common ethnicity is not being driven by a specific industry and is not particularly sensitive to product category divisions. Importantly, product categories that are unlikely to be affected by historic preferences (such as chemicals, manufactures, and machinery) are positive, significant, and in line with the main results.

This indicates that preferences are unlikely to be a driving factor in the results and demonstrates that the instrument is not being influenced by historically persistent preferences. If it were, the effect would be driven by product categories such as agriculture, raw materials, or organic oils—all products that could conceivably have been present in precolonial trade systems—and the results for these categories would be consistently larger than the other product categories.

[22] This division of categories combines manufactures and miscellaneous manufactured articles (categories 6 and 8 in the SITC system) for simplicity. Keeping these categories separate does not change the results.

Figure 1.4. Product-specific Coefficients



Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network.

The results from the inclusion of controls for factors such as infrastructure, institutions, and endowments combined with the evidence provided for preferences help to confirm the historical evidence and the discussion in Section 3.3. While these are all factors that could influence the relationship between the instrument and modern trade, the results presented here demonstrate that the instrument is robust to the inclusion of these other factors. This evidence contributes to satisfying the exclusion restriction, indicating that the effect detected in the main results is the effect of common ethnic groups on trade between countries.

B.2 Alternative Instrument

Along with sensitivity to the construction of the modern common ethnicity variables, the results may also be sensitive to different formulations of the instrument. I provide a test of the sensitivity of the results to the instrument used by constructing an alternative variable for historical trade networks. Table 1.12 presents the results of using a weighted shared network variable.

Whereas the instrument used in the main analysis is an unweighted count of shared trade posts, the variable used here incorporates a weighting system for each post based on an approximation of its size. While the size of specific trade posts is not mentioned in the sources, this variable uses the number of times each post is mentioned in the sources as a proxy for size. This rests on the assumption that larger trade posts would be more important within the networks and therefore more frequently mentioned in the sources.

The construction of this instrument is identical to that in the main analysis except for the inclusion of the various weights:

$$H_{ij} = \sum_n (\delta_{in} + \delta_{jn}) R_n$$
$$\text{with } \delta_{in} = \sum_p \omega_{ip}$$
$$\text{and } R_n = \begin{cases} 1 & \text{if } \delta_{in} > 0 \text{ and } \delta_{jn} > 0; \\ 0 & \text{otherwise.} \end{cases}$$

where δ_{in} is the number of weighted trade posts in country i belonging to network n given by the sum of the number of mentions, ω_p, across all trade posts p. Similarly to the original instrument, this variable measures how integrated countries are within shared networks. The inclusion of the weights gives more prominence to countries that contain trade posts that were more important within the networks.

The results presented in Table 1.12 are in line with the main results in Table 1.5. The incorporation of this construction of the instrument does not impact the results and demonstrates the robustness of the historical data as an instrument for common ethnic links.

Table 1.12. Alternative Instrumental Variable

	(1) OLS	(2) IV	(3) Reduced Form
Ethnic Link	0.495**	0.863**	
	(0.163)	(0.325)	
Shared Network			0.022**
			(0.007)
CYFE	Yes	Yes	Yes
N	16523	16523	16523
R^2	0.451	0.690	0.451

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is a dummy variable; Shared Network is the number of shared network trade posts; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

B.3 Religion

The sometimes strong connections between religion and ethnic identity could lead to a difficulty in isolating the ethnic effect from any effect caused by similar religious affiliation.

To address this, I develop several measures of religious similarity that follow a similar construction to the ethnic variables used in the analysis.

The first measure is similar to the Ethnic Link variable presented in Section 2.1. It captures whether the countries share a majority religion:

$$Religion_{ij} = \begin{cases} 1 & \text{if } \exists\ d \text{ for which } \zeta_{id} > 0.5 \text{ and } \zeta_{jd} > 0.5; \\ 0 & \text{otherwise.} \end{cases}$$

where ζ_{id} is the percentage of the population in country i of religious denomination d. A majority threshold was chosen due to the limited number of religions and the presence in all countries of members of at least one shared religion. With a majority threshold it also incorporates the potential commercial and legal power of the religion's adherents.

The second measure is an index of religion analogous to the EPI of Section 2.1. It captures the sum of shared proportions of religions within the countries. This measure is not dependent on population thresholds:

$$RPI_{ij} = \sum_{d} \zeta_{id}\zeta_{jd}$$

where ζ_{id} is the percentage of the population in country i of religious denomination d.

The third measure is an index of religious similarity between the countries (following the methodology of the ECI in Section 2.1):

$$RCI_{ij} = 1 - \frac{\sum_{d} |\zeta_{id} - \zeta_{jd}|}{2}$$

where ζ_{id} is again the percentage of the population in country i of religious denomination d. The ζ terms are subtracted across the two countries and then summed across all religions in the country pair. The index ranges from 0 to 1, with a larger value indicating more similarity.

The first column of Table 1.13 presents the results of including the common majority religion variable. The second column and third column show the results of including the two indices, the shared proportion index and the religious similarity index respectively. The main results for a common ethnicity remain unchanged through the inclusion of all three religious variables.

Islam Within the wider religious milieu of Africa, Islam was historically prominent across trade networks and among trading populations. This means that it could have an outsized influence on the results and that the instrument may be capturing the effect of networks based on Islam specifically rather than ethnicities. To account for this, I develop three variables that measure the Muslim population of countries identically to those constructed for religion more generally.

The construction of each variable is the same as those previously, except that they measure the proportions of Muslim populations only rather than being summed across all

Table 1.13. Common Religion Controls

	(1) Majority Religion	(2) RPI	(3) RCI
Ethnic Link	0.869**	0.879**	0.876**
	(0.323)	(0.322)	(0.324)
CYFE	Yes	Yes	Yes
N	16523	16523	16523
R^2	0.690	0.690	0.690

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network; Religious data comes from Maoz and Henderson (2013); ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

religions. The results are presented in Table 1.14. As with the results for religion more generally, the inclusion of variables controlling for Muslim populations does not affect the results in a significant way.

Table 1.14. Islam Controls

	(1) Majority Religion	(2) RPI	(3) RCI
Ethnic Link	0.882**	0.891**	0.885**
	(0.324)	(0.321)	(0.324)
CYFE	Yes	Yes	Yes
N	16523	16523	16523
R^2	0.690	0.690	0.690

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network; Religious data comes from Maoz and Henderson (2013); ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

B.4 Ethnic Variable Construction

The fluid nature of ethnicity and potential subjectivity of ethnic groupings presents a problem for the definition of the ethnic variables used in this chapter. To confront this, I test the sensitivity of the results to variations in both population thresholds and ethnic group definitions. The results of three specifications making these adjustments are presented in Table 1.15.

The first two columns present results for variations in the population threshold. In the benchmark specification the existence of any population, regardless of how small, is used as the threshold for the presence of an ethnic group. In the first column, this threshold is raised to 1000 people. The results produce a lower magnitude for the coefficient, but it can not be said to be significantly different from the value in the specification with the lower threshold. Similarly, raising the population threshold to 10000 people in the second column does not change the results.

The definition of ethnicity used in this chapter is one that favours self-identification of the groups themselves, potentially leading to more splits between groups (and therefore more individual groups). The specification in the third column changes the definition of an ethnic group to a broader view of what constitutes an ethnicity. This clumps previously individual ethnicities into larger groups, reducing the number of ethnic groups in the data to 2209. Despite this broader view of what constitutes a unique ethnic group, the effect of a common ethnicity on trade is not significantly different from the baseline specification in Table 1.5.

Table 1.15. Alternative Ethnicity Specifications

	(1) Pop=1000	(2) Pop=10000	(3) Broad Groups
Ethnic Link	0.649**	0.837**	0.711**
	(0.321)	(0.329)	(0.289)
CYFE	Yes	Yes	Yes
N	16063	15104	14654
R^2	0.697	0.696	0.663

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

B.5 Alternative Distance Measures

The gravity model relies on distance as a variable to capture generalised transport costs between countries. This can make the results sensitive to the measure of distance used, particularly given the variety of ways to measure distance between two countries. The measure used in this chapter is the direct great-circle distance as measured between the capitals of the countries. In order to test the robustness of the results to variations in this measure, three other distance variables are also constructed. The results of these are presented in Table 1.16.

Column 1 includes the first of these: the direct great-circle distance between the largest cities of each country. This captures the distance between the main economic hubs of the countries, as opposed to the political capitals which may not be the centres of commerce. The use of this measure does not change the results for the coefficient on the Ethnic Link variable.

Table 1.16. Alternative Distance Measures

	(1) Major City	(2) Centroid	(3) Transport
Ethnic Link	0.874**	0.563**	0.671**
	(0.324)	(0.279)	(0.283)
CYFE	Yes	Yes	Yes
N	16523	16523	16523
R^2	0.690	0.692	0.695

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

The second column presents the results for centroid-based distance. This distance is calculated as the great-circle distance between the geographic centroids of the countries. This accounts for the fact that capitals and major cities may not be located near the middle of countries and so the distance may not capture the distance needing to be travelled to reach a border, this is particularly acute in large countries and countries that are oriented along a particular axis. The use of this distance measure reduces the coefficient on Ethnic Link, but it remains strongly positive and significant. Despite the reduction in value it cannot be statistically distinguished from the main results in Table 1.5.

The last column displays the results for a more complex, transport-based measure of distance. This distance uses data from the CERDI database which consists of raw distance data for three relevant measures: 1) the distance between the two countries' relevant ports; 2) the distance between the port and capital of each country; 3) the land distance between the capitals of the two relevant countries. In cases of landlocked countries that do not possess their own territorial port, the foreign port that is most used as an entrepôt by the specific country is identified and distances are calculated using it.

The distance between the two relevant ports of a bilateral pair is calculated in kilometres using regularly used sea lanes for shipping cargo. The distance between port and capital for each individual country is determined in kilometres using the most direct road route between the two points. These three distances (the distance between ports and the distance between port and capital for the two countries in a given pair) are then summed to achieve the total surface travel distance for the trading pair. In cases where the land distance (calculated as the distance between the two capitals of the relevant countries via major road systems within each country) between two countries is shorter than the sum of the two capital-port distances, it is assumed that an over-land route would be primarily used for trade and this distance is used.

The inclusion of this measure for distance also reduces the magnitude of the coefficient for a common Ethnic Link, but it remains positive and significant. Again, despite the reduction in magnitude, it is not statistically different from the main results.

B.6 Zero Trade and Missing Data

Trade flow data can present a significant number of zero flows or missing values. Due to the logarithmic transformation of the gravity model, these data points are excluded from the standard analysis. This can introduce selection bias in the results when only strictly positive values are analysed. To address this issue of bias Silva and Tenreyro (2006) develop a non-linear estimation technique using the Poisson pseudo-maximum likelihood method (PPML). With this technique, the model is estimated in an exponential form so that trade flows between countries can be estimated in levels, maintaining the zero flow values. This modified estimation equation is given by:

$$
\begin{aligned}
E(X_{ijt}) = exp\{&\beta_0 + \alpha \ln(Y_{it}Y_{jt}) + \beta_1 \ln(d_{ij}) \\
&+ \beta_2 land_{ij} + \beta_4 isle_{ij} + \beta_5 bord_{ij} \\
&+ \gamma_1 ethn_{ij} + \gamma_2 lang_{ij} + \gamma_3 col_{ij} \\
&+ \theta_1 curr_{ijt} + \theta_2 cust_{ijt} + \theta_3 rta_{ijt} + \theta_4 grp_{ijt} \\
&+ \mu_i + \mu_j + \nu_t + \epsilon_{ijt}\}
\end{aligned}
$$

The dataset used in this chapter has very few zero values for the trade variable (22 cases or 0.009% of total available observations), but the number of missing values is higher (8150 cases or 31.6% of total available observations). In order to check the robustness of the original OLS estimation, I use two variations of the PPML model. In the first, I use the traditional method to include the zero-value cases in the estimation. The second variation replaces all missing values for the trade variable with zeros and re-estimates the PPML equation. This is done to account for the possibility that missing values in the data could merely be misinterpreted zeros. The results of these two estimations are presented in Table 1.17 along with the original linearised model for comparison. In both cases the coefficients are slightly larger indicating that rather than over-estimating the effect, as is more common in gravity models, the OLS specification appears to be under-estimating it. This seems to indicate that countries without shared ethnicities tend to not only trade less, but also tend to not trade.

B.7 Regional Sensitivity

Due to the geographic nature of ethnic distributions, the sensitivity of the results to regional variations in the sample is an important element. It is possible that some regions have specific ethnic groups that are large and widely distributed across multiple countries,

Table 1.17. PPML Results

	(1) OLS	(2) PPML	(3) PPML+
Ethnic Link	0.495**	0.661**	0.705**
	(0.163)	(0.187)	(0.189)
CYFE	Yes	Yes	Yes
N	16523	16893	22970
R^2	0.451	0.808	0.810

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is a dummy variable; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

potentially skewing the results. To explore this I drop some regions from the sample that have unique features with respect to ethnic distributions and so may be disproportionately influencing the effects. These results are presented in Table 1.18.

The first column of this table presents the results just for sub-Saharan Africa, excluding North African countries from the sample. The second column expands this to also exclude Sahelian countries from the sample. These two regions (North Africa and the Sahel) have a number of ethnic groups that are widely distributed across many countries within the regions (such as the Tuareg) and excluding them tests the sensitivity of the results to the presence of these types of ethnic distributions. As the first two columns show, the coefficients are not significantly different from the baseline estimation. The third column excludes West Africa, a region that has many different ethnic groups that are frequently shared between smaller countries. Excluding this region tests the sensitivity of the results to this type of frequent sharing of ethnic groups. The results in this case also remain unchanged from the baseline. Finally, Southern Africa was excluded as it also has fewer ethnicities that tend to be distributed among larger countries but it lacks the geographic element of the Sahara. Though the coefficient is larger in magnitude, it is not found to be significantly different from the baseline estimation.

B.8 Variations in Sample Period

I also test the sensitivity of the results to variations in the time period of the sample. Figure 1.5 presents the coefficients for year cross-sections of the data estimated independently. The coefficients tend to be larger towards the end of the sample, but they fluctuate around the baseline coefficient estimate (the blue line). An exception to this is the coefficient for the year 2001 which is anomalously negative (despite not being significant). An explanation for this discrepancy could be related to the sensitivity of informal networks to fluctuations in global trade trends. Global trade in 2001 experienced a sharp decline across most countries and it is possible that this effect could be more pronounced

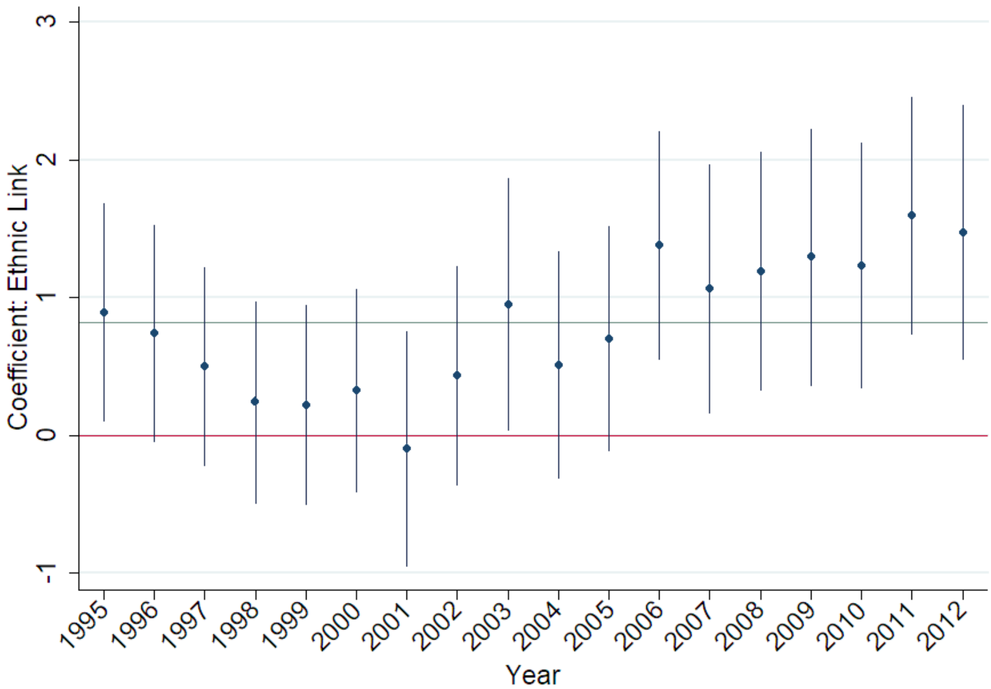
Table 1.18. Regional Robustness Checks

	(1) SSA	(2) SSA -Sahel	(3) -West	(4) -South
Ethnic Link	0.713** (0.337)	0.965** (0.363)	0.860* (0.467)	1.032** (0.350)
CYFE	Yes	Yes	Yes	Yes
N	11210	8573	8529	10905
R^2	0.666	0.697	0.699	0.698

Standard errors in parentheses, clustered at the country-pair level
* Significant at 0.10; ** Significant at 0.05
Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

between countries that rely on informal networks (such as ethnic networks) to promote trade rather than more formal institutional relationships that would be able to weather a downturn more effectively.

Figure 1.5. Year Cross-section Coefficients



Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network.

I supplement this with broader cross-sectional estimations based on averages over multiple years in the sample. These estimations average the time-varying variables over the given sub-period of years and estimate the IV model as a cross-section. Table 1.19 presents the results for two different periods. The first column averages across the entire sample (1995-2012) while the second column contains results for the final 10 years of the sample period (2003-2012). The coefficients in both cases are positive and significant. Though they are larger in magnitude than those in the main estimation, they are not found to be significantly different from the baseline results.

Table 1.19. Year Average Cross-sections

	(1) 1995-2012	(2) 2003-2012
Ethnic Link	0.991**	1.118**
	(0.375)	(0.397)
CFE	Yes	Yes
N	1207	1175
R^2	0.749	0.741

Standard errors in parentheses, clustered at the country-pair level

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total bilateral trade; Ethnic Link is instrumented by Shared Network; ln(Distance), Landlocked, Island, Border, Language, Colony, Currency, Customs, RTA, and Group are all included.

C Appendix: Historical Trade Networks

C.1 Historical Background

The idea that cultural groups can facilitate trade through intra-cultural networks in Africa has a long historical tradition. There is much anecdotal evidence in the literature on the commercial dominance of various ethnic groups in certain places and at certain times. The networks created by these groups have been attested through the early modern period, but also pre-date colonialism. This historical literature broadly focuses on five commercial areas which can be collapsed into three general regions: 1) North Africa, West Africa, and the Sahel; 2) Central Africa (which is sub-divided into: a) the Congo Basin, b) the East African Coast, and c) South-East Africa); and 3) the Horn of Africa. Below are brief historical sketches of trade and commerce in each region.[23] These regions are created organically by the data on trade posts and the connections between them. They are effectively five semi-isolated clusters of trade posts that are connected to each other and rarely (if at all) to posts in other clusters.

C.1.1 North Africa, West Africa, and the Sahel The various networks that link together in the northern half of Africa are some of the best attested networks on the continent. Records of trade routes linking the North African coast to the Sahel through the Sahara date back to ancient times. Documented links between these networks and the West African coast go back nearly as far. Throughout the medieval and early modern periods, cities in these areas maintained strong trade networks with each other for goods of all kinds. While the most commonly attested forms of trade were the export of gold, slaves, and salt for manufactured and luxury goods, this is primarily due to that part of trade being of most interest to those documenting it. Despite this cash-crop bias, trade in all kinds of other goods are documented by both local sources and outside observers. The value of these trade routes and the prosperity they created for the cities along them led to the rise of many empires and powerful states. In the west, the Malian empire and later the Songhai empire rose to dominate the Niger valley while in the east Kanem-Bornu emerged to control the area around Lake Chad. Between these extremities the various Hausa kingdoms expanded as major ports of call for the trans-Saharan trade. On the West African coast many cities and kingdoms exerted power through monopolisation of export goods or dominance of major trade posts. The various Akan kingdoms derived their power from dominance in the gold trade, while the cities along the coast and in the Niger delta gained power by serving as major entrepôts for early European traders.

C.1.2 Central Africa Trade networks in the region of Central Africa are more complex in their interaction than in either of the other two regions. Networks in this region

[23] For a more complete discussion of historical linkages between cities in these regions see Austen (2010), Bovill (1995), Chirikure (2017), Coquery-Vidrovitch (2005), Gray and Birmingham (1970), King and van Zwanenberg (1975), and Ogot (1974).

can be subdivided into three areas: a) the Congo Basin, b) the East African Coast, and c) South-East Africa. The difficulty derives from the fact that while trade routes and networks within these three areas are relatively well attested and documented in the pre-colonial period the links between the areas are not. Often, the links that do exist are between cities in two of the three areas which heavily imply connections to the third (A is connected to B in one source; B is connected to C in a different source; but the link between A and C directly is missing). Similarly, there are many more cities within each area that are attested to in isolation of the wider region than there are cities included in networks between the areas (and therefore linking the region together). Conversely, there are cities that are discussed solely in their relation to the wider regional links and not in terms of any connection to local area networks. Below, are brief descriptions of each area in isolation.

a) Congo Basin Trade networks in the Congo basin area were primarily focused on trade between the interior and the coast. This was largely driven by the export of metals, ivory, and slaves for manufactured goods and luxury items. Resources from the interior and the wider Congolese watershed moved up and down the river system to be traded at the coastal ports. While trade is often attested as being oriented towards western coastal cities, there is evidence of trade with the eastern coast and the south-eastern networks. The wealth of these networks gave rise to numerous cities and settlements that participated either in the extraction of raw materials or as way-stations along the sometimes circuitous routes. Kingdoms such as Kongo in the west and Luba in the east profited from their positions along these routes and their power expanded further after the arrival of European traders. Other kingdoms, such as the Lunda Empire in the south-central Congo basin benefited through extraction of resources for trade and rich deposits of metals and other goods.

b) East African Coast The East African coast has been a principal trade hub since antiquity. It was called Azania in the Greco-Roman period and has been attested to in numerous sources throughout history. The eventual rise of Swahili city-states on the coast and settlements on offshore islands was a direct result of the prosperity created by the trade networks of the Indian Ocean. These cities would serve as points of exchange, buying goods from the interior and then selling them onwards to their networks in the Indian Ocean. Raw materials such as ivory, spices, timber, and gold would make their way from these ports into the Indian Ocean in return for finished goods such as silk, porcelain, and cotton textiles but also other raw materials like coffee, pepper, and sandalwood. As a result of this trade the cities of Bagamoyo, Kilwa, Mzizima, Malindi, Mombasa, Pangani, and Tanga became some of the richest cities in the region. While kingdoms such as Buganda, Rwanda, Burundi, and Maravi centred on agriculture and raw materials rose in the interior.

c) South-East Africa South-East African trade networks grew around trade in ivory and precious metals, primarily gold from the southern parts of the area and copper from the northern parts. This trade also focused on exporting raw materials for manufactures and luxury items, creating incredibly wealthy kingdoms. Control of lucrative gold fields led to the rise of kingdoms based on its extraction and trade such as Mutapa and Butua, while further to the south-west the Mapungubwe and later Zimbabwe kingdoms controlled the trade in ivory and various metals. While this area was largely linked into coastal trade with the Indian Ocean, routes and networks attested to mention primarily southern coastal ports, separating it from the primary activity on the East African coast. Despite this, there are documented links with the more northern sections of coast and the deeper interior of the Congo basin.

C.1.3 The Horn of Africa While not as significant a region as the others mentioned, the Horn of Africa is considered separately from Northern Africa more generally due to its unique position and orientation. While the networks in the north of Africa were largely oriented towards the Mediterranean or the Atlantic coast, the networks in this region were, since ancient times, heavily connected to the Arabian peninsula and the Indian subcontinent. Since the Iron Age, what are now northern Somalia, Djibouti, and Eritrea were a significant source of valuable incense and spices and the commercial and political connections to the southern tip of the Arabian peninsula date to this era. In this sense, the Horn of Africa's orientation was distinctly different from that of the majority of pre-colonial trade in the northern half of Africa; but it was also distinct from the networks of Central Africa with many of its trade routes having only weakly implied connections to the wider Central African networks until the arrival of the Portuguese, and later the Omanis.

C.2 Instrumental Variable Construction

Constructing a dataset of pre-colonial[24] trade centres and routes in Africa requires the accumulation of relevant data from a disparate set of sources. This naturally runs into problems given that different sources will emphasise different centres within the various areas and different routes depending on the purpose of the document's narrative. This is a more significant challenge for trade centres that are more minor or regional in nature and not connected to larger networks, therefore being less likely to feature in discussions of trade in wider regions. In order to overcome this problem a diverse set of sources dealing with a range of issues and locations is used in an effort to obtain a multidimensional perspective of trade centres and routes on the continent.[25]

In order to assemble a meaningful dataset for the instrument, each settlement (hereafter referred to as trade posts) is identified and assigned a set of geographic coordinates

[24] The information used spans the medieval period up until the early nineteenth century.

[25] Sources included Bovill (1995), Chirikure (2017), Coquery-Vidrovitch (2005), Gray and Birmingham (1970), King and van Zwanenberg (1975), and Ogot (1974).

corresponding to its location. The routes connecting each trade post to another trade post are less helpful in the broader construction of the instrument beyond being indicators for which network a given trade post is a part of. The reason for this is two fold: first, the accuracy with which the exact route between any two trade posts is recorded is highly suspect and conditioned on the attentiveness of the recorder, the time it was recorded (as routes can change), and other idiosyncratic factors; and second, the idea that there existed only one route between any two trade posts is an unjustified oversimplification of how trade would flow between any two sites. For these reasons, the routes between any two sites are not used directly in the analysis and are presented as straight lines (in Figure 1.1) for clarity of illustrating the various networks.

For a given trade post to be considered part of a larger network, it must be connected in some way to that network. This is how the five area networks, and three regional networks, are constructed and where the problems regarding the Central African region identified previously emerge. The difficulty with the scarcity of attested connections between the three areas within the Central African region means that there are three ways in which a trade post can be classified within the region: 1) the trade post can be mentioned only in reference to its area; 2) the trade post can be mentioned only as a member of the wider regional network; and, 3) the trade post can be indicated to be a part of both an area network and the wider regional network. The approach chosen to deal with this issue is the path of least assumptions. The network membership of each trade post is accepted at face value with as little interpolation of wider linkages as possible. This creates a situation where the three areas are each themselves considered a network and the wider region is considered an additional network. Trade posts are able to be members of either or both of their area network and the wider regional network. This creates a situation where there are in fact six total networks within the data, corresponding to the regions of North Africa, the Horn of Africa, the areas of the Congo Basin, the East African Coast, South-East Africa, and the wider region of Central Africa. Trade posts classified as members of only their area are not included in the Central African regional network and trade posts that are classified as only members of the Central African regional network are not included in any potential area network. Trade posts that overlap the two networks are included in both their assigned area network and the regional network.

D Appendix: Data Information

Data pertaining to ethnicity are always difficult to quantify, and Africa is no exception to this. The difficulty exists primarily in the very subjective nature of ethnic groups, and nowhere are the boundaries of ethnic identification more fluid than the continent of Africa. For example: the Nyanje of Malawi speak ChiChewa and are usually considered to be a sub-group of the Chewa ethnic group, but many researchers classify them as a separate group. Similarly the Baoulé of Ivory Coast fall under the umbrella of the Akan people both linguistically and ethnically; however, the Asante, another sub-group of Akan people, could be considered separate from the coastal Akan peoples in Ghana. This fluidity of division from country to country poses many problems in terms of ethnic classification between and within countries. This appendix details the dataset and data collection process for the ethnic data used in this paper.

D.1 Data Composition

Due to the fluidity and inconsistency inherent in many ethnic classifications, ethnic data for countries in Africa can be difficult to come by. Often domestic census data does not contain ethnic information; and when it does, the information can be imprecise or incomplete. A single source of ethnic data for the countries concerned is necessary to avoid problems that could come from using different and unreconciled data sources, including differing degrees of accuracy. Fortunately, missionary groups within the continent track multitudes of data, particularly on conversion rates to their chosen religion. One group (The Joshua Project)[26] possesses two qualities which make its data useful for the purposes of this study: it is pan-African in its scope and it maintains data on all groups of people (not just those it has converted).

This organisation has compiled data on ethnic groups in every country of the world, combining data from numerous sources to create a complete record of the population levels of each ethnic group for each country. The database uses national census data as a base and supplements this with significant input from local in-country experts and researchers. This takes advantage of a uniquely local understanding of ethnic groups and ethnic divisions within countries and provides a level of country-specific detail to the data that other datasets may lack. The data are cross-checked with larger ethnolinguistic datasets such as Ethnologue, the World Factbook, and UN sources to verify and correct any inconsistencies in the locally gathered information.

The data come standardised at a rough level and are formatted to be consistent across countries. The raw dataset errs on the side of splitting groups rather than combining them, which generally increases the representativeness of the data. By including sub-groups as potentially separate ethnic groups, there is less chance of any individual group

[26] More information about the raw data can be found at https://joshuaproject.net.

being missed or subsumed into a larger group. The population figures in the dataset are current for the year 2012.

D.2 Definition of Groups

For the purposes of this chapter the raw data need to be rendered usable in a cross-country comparison of ethnic groups, leading to a more refined and standardised dataset. The granular structure of the data includes all self-identified groups in a given country and the total population within each of these groups. The propensity for the raw data to split and uniquely identify local groups means that groups that are shared between countries are not always identified as such due to differing local names and identifiers. There are three general cases of this type of splitting: different local names, different language identifiers, and different sub-group identification.

Local Names The first source of splitting is the use of different local names in different jurisdictions for the same ethnic group. This is a common practice, particularly for groups that are present in many countries with local languages (and different national languages) creating different terms for the same group. The Fulbe people are a prime example of this, with different names in different countries based on both local languages and national languages. Former French colonies tend to categorise them using the the French term 'Peul', while English-speaking countries tend to use the terms 'Fulani' or 'Fulbe'. Adding complexity, different local languages also use different terms with Bambara-speaking jurisdictions using the term 'Fulaw' and Wolof-speaking areas sometimes using the name 'Pël'. These differences in local names do not signify differences in the ethnic groups they describe and so are reconciled across the different jurisdictions in order to standardise the ethnic composition of countries.

Language Identifiers A second splitting feature is the division of ethnic groups along linguistic lines. This is where ethnic groups that are largely homogenous are split due to differences in local dialect or *lingua franca*. An example of this are the Hutu people which are divided between the mutually intelligible dialects of Kirundi and Kinyarwanda. This division and other divisions along English, French, or Portuguese lines are linguistic divisions overlaid on ethnic categories. Since the focus of this chapter is ethnic group connections rather than linguistic group connections, these linguistic splits are grouped together into the common ethnic definition.

Sub-group Identification The final type of splitting is the selective use and reporting of ethnic sub-groups between jurisdictions. This is where different countries or areas use different levels of identification for ethnic groups with some jurisdictions splitting relatively broad groups into more specific sub-groups and other jurisdictions using the broader categories. For example, the Akan group in West Africa is split in different ways between different countries (with sub-groups such as the Fante and Akwamu reported in some countries and not others). For the purposes of this chapter, this type of splitting is accommodated through the grouping of the various sub-groups within the broader ethnic group.

In terms of reconciling these splits and handling the cases of overlapping ethnic identities or sub-groups of larger populations I have employed whichever grouping is considered standard, balanced with the commonly recorded impressions of the groups themselves. Often, this amounts to whether a sub-group is large enough to be considered distinctly separate. This has resulted in situations such as the Baoulé not being included with the Asante within the Akan group, while the Nyanje are combined with the Chewa; and the various sub-groups (such as the Jelgoobe and Wodaabe, among many others) in West Africa being considered parts of the larger Fulbe group.

While this has led to some clumping of smaller groups and naturally increases the number of countries sharing one ethnic group, I believe the method employed best preserves the potential effect of an ethnic group on trade networks and is shown to be robust to different population thresholds and types of group clumping. This process resulted in the creation of 2381 different ethnic classifications for the continent from an original set of approximately 3700; it is done with an eye to maintaining self-identification within groups, as this self-identification and understanding of commonality at the micro-level is the primary pathway through which an ethnic network would be forged.

D.3 Potential Measurement Error

The subjectivity of ethnic data and the judgements made in compiling it introduce the potential for several sources of measurement error. The first of these sources is the structure of the core dataset. The difficulty in combining several sources of ethnic group population data presents the potential for inconsistencies across the groups recorded as discussed in the previous subsection. The efforts taken in this chapter to harmonise these issues help to ameliorate this problem to some extent, but there remains the risk that certain groups will be miscounted or go uncounted.

The second source of potential measurement error is in the process undertaken to combine and divide the groups from the raw data to create consistency in ethnic group categorisation across countries. This process is necessarily arbitrary to some extent and there is no method that will be universally accepted. In this chapter I have sought to maintain local identity and create groups in line with commonly accepted and self-identified attachments. This does not make the process free from error.

Despite these risks, the data and the process to construct it is relatively robust to these issues. This is particularly true in terms of the potential problems around clumping and splitting, with the data being demonstrably robust to different group divisions. While the potential for measurement errors is relatively small in terms of shared ethnicities across countries due to the aggregation involved in constructing the variables, they cannot be completely discounted.

Table 1.20. List of Countries

Algeria	Guinea	South Africa
Angola	Guinea-Bissau	South Sudan
Benin	Kenya	Sudan
Botswana	Lesotho	Swaziland
Burkina Faso	Liberia	Tanzania
Burundi	Libya	Togo
Cameroon	Madagascar	Tunisia
Cape Verde	Malawi	Uganda
Central African Republic	Mali	Zambia
Chad	Mauritania	Zimbabwe
Comoros	Mauritius	
Congo, Democratic Republic	Morocco	
Congo, Republic	Mozambique	
Côte d'Ivoire	Namibia	
Djibouti	Niger	
Egypt	Nigeria	
Equatorial Guinea	Rwanda	
Eritrea	Sao Tome and Principe	
Ethiopia	Senegal	
Gabon	Seychelles	
Gambia	Sierra Leone	
Ghana	Somalia	

Chapter II

Interrupted Lines: Conflict, Trade Flows, and Transport Costs

1 Introduction

Transport costs are an important factor in determining trade flows, but they are frequently thought of as the cost of shipping goods from one country to another, including freight costs, transport time, and insurance. In practice however, transport costs also include many more indirect costs: administrative paper work, regulatory compliance, logistical considerations, network costs, information costs, and uncertainty. One important, yet often overlooked, aspect of these costs is the presence of conflict, particularly if the conflict lies along trade routes needed to access external markets. Conflict can have an important impact on trade costs, and consequently trade flows themselves. Landlocked countries in particular need access to external markets via trade routes through other countries and their ports. The case of Malawi in Figure 2.1 demonstrates this clearly. The threat of conflict along Malawi's lines of trade through Mozambique caused significant disruption to economic activity and trade during the Mozambican Civil War. Malawi, as a landlocked country, was particularly vulnerable to disruptions as it did not have alternative access to global markets through its own ports. For countries such as these, conflict along their trade lines can have a substantial impact on their ability to access markets for trade. Analysing these effects is the objective of this chapter.

To approach this issue, this chapter uses a modified version of the gravity model with aggregated trade data to test the impact of adjacent conflict on trade flows into and out of countries. Critically, I exploit the geographic elements of both country location and conflict location to identify the impact of conflict on trade routes. Adjacent conflict is potentially more disruptive for a landlocked country than a coastal country as it can disrupt trade routes and access to external markets. Additionally, adjacent conflict in a country that is coastal (shoreward) is more disruptive as it explicitly blocks access to external ports needed by landlocked countries to export and import goods.

I use the Major Episodes of Political Violence (MEPV) dataset to construct conflict variables for both direct and adjacent conflict effects. The results indicate that there is a strong negative effect of adjacent conflict on trade. This effect appears to be stronger for landlocked countries and in particular when the adjacent conflict is in a coastal country. This provides evidence for a supply-driven effect on trade in which the adjacent conflict is disrupting trade routes and the access of affected countries to international markets.

The identification strategy rests on disentangling the different impacts of adjacent conflict though differences in geography. I address concerns about time-varying omitted variables at the country level and at the wider regional level by incorporating variables to capture changes in institutions and trade agreements. While the baseline findings are re-enforced by the inclusion of these indicators, they are not robust to controlling for more general regional trends. The spatial nature of the identification strategy makes it difficult to fully distinguish the effects of adjacent conflict from these wider regional trends.

This chapter contributes to the literature by seeking to explicitly disentangle the different effects of adjacent conflict on trade and exploring the specific effect of adjacent conflict on trade routes and the ability of countries to access external markets. This is closest to the work of Marano et al. (2013) which looks at the production impacts of conflict in adjacent countries and Qureshi (2013) which explores the effect of adjacent conflict on trade from a demand-oriented perspective. This chapter supplements Marano et al. (2013) by expanding the scope of the effect of adjacent conflict to trade flows and transportation costs (after production). The primary focus of this chapter is the effect of adjacent conflict on trade routes and transport costs, complementing the work of Qureshi (2013) by adding a supply-focused perspective to the discussion of the effects of conflict on trade in neighbouring countries. In pursuing this, the results of this chapter present evidence for an impact of adjacent conflict on trade flows through a supply-oriented channel involving transportation costs and the disruption of trade routes.

Much of the existing research on the relationship between trade and conflict deals with the negative impact conflict has on trade flows when one of the trade partners experiences a conflict. This generally builds on the 'Democratic Peace' tradition and its extensions in political economy.[1] This has expanded into the economics literature with Goenner (2011) providing a bridge by investigating the endogenous nature of trade as a variable in a study on its effectiveness at reducing conflict. Nitsch and Schumacher (2004) and Mirza and Verdier (2008) have expanded this to include the negative effects of terrorism; while Bayer and Rupert (2004) find similar results for civil wars. Counter to the prevailing results are some studies (Garfinkel et al., 2008; Bah and Tapsoba, 2010) that look at possible short-run positive effects from direct conflict exposure.

Few papers expand beyond these direct effects into the impact of adjacent conflict on trade flows. Literature that does touch on the impact of conflict in adjacent countries tends to focus on other economic factors such as growth or income effects.[2] Parallel

[1] See Oneal and Russet (1997), Oneal and Russet (1999a), Oneal and Russet (1999b), and Hegre et al. (2010) among others.

[2] Ades and Chua (1997), Murdoch and Sandler (2002), Murdoch and Sandler (2004), Collier and Hoeffler

to this are investigations of more complex relationships between direct conflict and trade such as Kang and Reuveny (2001), Blomberg and Hess (2006) and Glick and Taylor (2010) which all extend the analysis of direct conflict on trade from bilateral pairs to multi-lateral clusters. Notable exceptions to this are Marano et al. (2013) and Qureshi (2013) which address the impact of adjacent conflict explicitly.

Connected to this work are the linked issues of trade diversion and trade creation. Most of this research deals with the effects of trade agreements and their potential to divert and create trade within and between trade blocs. Discussion of the issues of trade diversion and creation constitute a broad literature that dates back to the work of Viner (1950) and Balassa (1967).[3] This includes work on the effects of traditional trade agreements as in Clausing (2001) and Datta and Kouliavtsev (2009); as well as more comprehensive agreements such as the European common market covered by Haveman and Hummels (1998) and Kahouli and Maktouf (2015). Magee (2008) and Eicher et al. (2012) take this work further with the inclusion of counter-factual analysis and Bayesian techniques to improve the estimations. This chapter addresses some of these related issues from a different direction by focusing on the hindering of trade flows and the interruption of trade routes.

Transport costs are important factors in shaping trade flows between countries and are central to the effect of conflict on trade. Significant work has contributed to the robustness of gravity models and documented the effects of transport costs and transport technology.[4] First among these is Engel and Rogers (1996) which identifies a generalised 'border effect' in transporting goods to other countries. Of note are Hummels and Schaur (2013) and Feyrer (2019), which both look at the effect of progressively decreasing air transport costs on bilateral trade, while Pascali (2017) investigates the gains in trade made due to the introduction of steam vessels during the nineteenth century. In the other direction, Feyrer (2009) looks at the impact of the closure of the Suez Canal during the 1967 Six-Day War on oceanic trade. This literature has even been extended further back to trade in the ancient world with Barjamovic et al. (2019). I build on the effects in these papers by adding adjacent conflict as a contributing factor to changes in transportation costs experienced by countries.

This chapter is divided into six sections. The following section provides a discussion on the mechanisms by which conflict impacts trade. Section 3 outlines the empirical approach and the data is covered in Section 4. Section 5 presents the results while Section 6 concludes.

(2004), and De Groot (2010) are all examples of this trend.

[3] Panagariya (2000) provides a comprehensive survey of the literature.

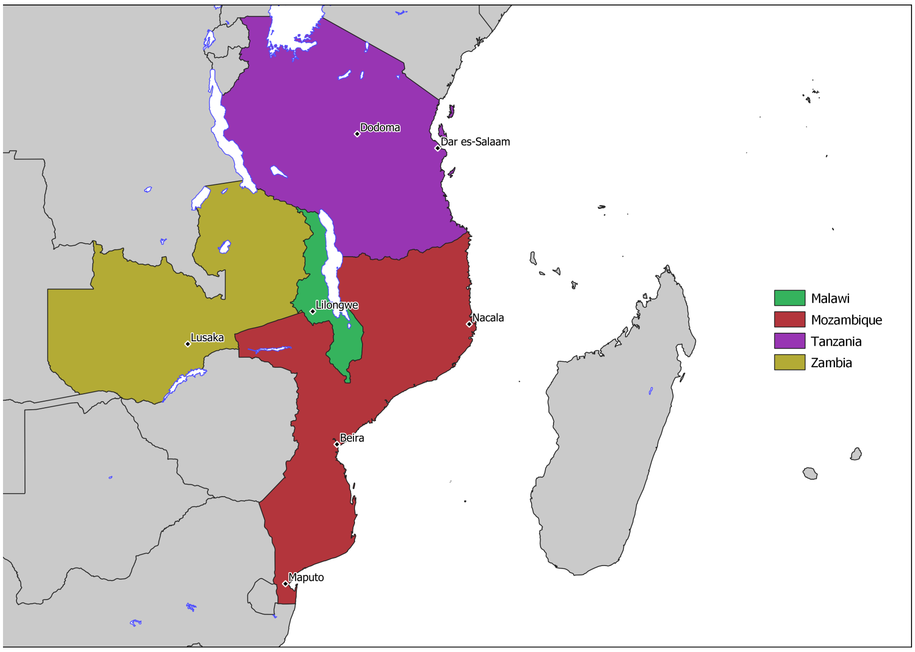
[4] For a more extensive survey of literature on gravity models and transport costs, see Head and Mayer (2014).

2 Mechanisms

Conflict in adjacent countries can impact the trade of a given country through two different channels: supply and demand. The supply channel is connected to market access and the ability to move goods in and out of the country whereas the demand channel is connected to the availability of markets for the goods traded. This section discusses these two channels in more detail and a methodology for disentangling the two effects.

Supply The supply effects of adjacent conflict relate primarily to the disruption of the flow of goods or increased costs associated with transporting them. The first of these involves interruption of trade routes through damage or threat to infrastructure needed to transport goods into and out of the country. This can include minor impacts on shipping at ports near the border and more major impacts on road and rail networks that travel through the conflict zone to reach a port. This can directly disrupt trade by preventing goods from travelling in or out of the country, but it can also increase trade costs by forcing importers and exporters to use more expensive, second best routes and methods of moving goods increasing trade costs and decreasing total trade flows. Additionally, the presence of a conflict zone may not always physically disrupt infrastructure or block trade routes, it can also increase the cost of transporting goods through the conflict zone. This includes costs such as increased insurance, stronger security, or more difficult logistical challenges.

Figure 2.1. Mozambican Civil War, South-East Africa



Note: Countries of south-eastern Africa, major cities of relevant countries are indicated.

Supply effects of this nature are common anecdotally, with the case of Malawi during the Mozambican Civil War (1977-1992) being the most prominent. As shown in Figure 2.1, Malawi is a relatively small landlocked country to the west of Mozambique which contains the ports of Beira and Nacala. These ports were, prior to the conflict, the principal entrepôts for Malawi with important rail and road links connecting it to these ports. During the civil war in Mozambique, these links were severed—in some cases literally, with rail lines being ripped up by some of the combatants. This led to a significant disruption of the ability of both Malawian exports to access foreign markets and imports to reach the country through Beira and Nacala. This scenario, while possibly an extreme case, is indicative of the potential supply-based problems adjacent conflict can cause.

Demand The demand effects of conflict in adjacent countries are related to a reduction in external demand for goods impacted by the adjacent conflict. This can either be directly related to the country being near a conflict zone or indirectly through substitution effects. The first of these involves a reduction in demand for goods from a given country because of its association with a conflict zone.[5] This can be interpreted as a 'blood diamond' effect where the goods from countries near or in conflict zones are considered tainted by the conflict, either because of perceptions that revenues from industries are fuelling conflict or because the industries are associated with combatants directly.

In addition to this direct demand effect, there are also more indirect substitution effects. These come from firms outside the conflict zone substituting away from goods from the country impacted by the adjacent conflict in search of potentially more reliable sources that would not be impacted by conflict-related disruptions. This pre-emptive shift depresses demand for goods from countries in proximity to the conflict zone and reduces trade.

In opposition to this is the potential for the substitution effect to work in the opposite direction, with trade in neighbouring countries increasing as firms outside the conflict zone substitute away from the country experiencing direct conflict. This potential would be increased if the neighbouring countries produce similar goods and are less affected by the supply-driven effects disrupting their trade.

These demand-driven effects are also common anecdotally with the most notable impacts being on diamonds and other minerals from conflict zones during the civil wars in Liberia, Sierra Leone, and Angola. These all involved concerted international campaigns to reduce the desirability of products from these regions in order to reduce the revenues going to combatants engaged in their production.

Separating the Effects The supply effect is driven by the combined geography of preferred trade routes and adjacent conflict. The ideal approach would combine geography, transportation costs, and conflict incidence to compute transportation costs adjusted for the presence of conflict. The complexity of this approach makes its use difficult with aggregate data. A simpler, and more direct, test involves identifying the disrupting effect

[5] This channel is what is primarily explored in Qureshi (2013).

of conflicts from the most apparent geographic cases: coastal countries and landlocked countries.

Consider a hypothetical case of three countries that all neighbour each other. The first country is a coastal country and suffers from an internal conflict. The second country is landlocked and requires access to the first country's ports in order to engage in trade. The third country is coastal, and so has its own ports. The conflict in the first country will affect trade for the other two countries. It is assumed that both of these countries will experience demand-driven effects of the adjacent conflict. However, the effect on the second (landlocked) country will be more severe due to the supply-driven effect because it requires access to the first country's ports. This difference in the total effect of adjacent conflict between the second and third countries is the supply-driven effect.

As this hypothetical example shows, separating the demand effect from the supply effect is contingent on two aspects of the geography of the three countries. The first is the location of the country not in conflict: the difference between being coastal and landlocked. The second is the relationship between the landlocked country and the country in conflict: the country in conflict must be shoreward to the landlocked country in order for it to have ports that the landlocked country needs access to. If the conflict country were also landlocked (inland of the other countries) then the adjacent conflict may not produce the difference in effects in the two other countries.

This can also be illustrated concretely with the anecdotal example shown in Figure 2.1. Tanzania is a country along the coast, to the north of Mozambique, with three oceanic ports—the largest by far being Dar es-Salaam. This means that Tanzania, unlike Malawi, was not reliant on accessing Mozambican ports in order to export or import goods and so was insulated from the conflict-based disruption to its trade of the Mozambican Civil War. Similarly, a hypothetical conflict in Zambia (which borders Malawi but is not a coastal nation itself) may have less impact on Malawian exports and imports than the conflict in Mozambique, since such a conflict would not restrict access to ports and external markets in the same way.

This chapter tests the hypothesis that adjacent conflict will negatively impact trade for a country through the disruption of trade routes by exploiting geographic variation. As outlined in the following section, by incorporating the location of the country (landlocked or coastal) and the location of the conflict (shoreward or inland) the impact of adjacent conflict through the supply channel can be isolated from other, demand-driven, impacts.

3 Empirical Approach

3.1 Gravity Framework

The goal of this chapter is to analyse the effect of adjacent conflict on a country's trade. There is an extensive literature using gravity models to analyse the relationship between trade flows and transportation costs.[6] Anderson and van Wincoop (2003) and Feenstra (2004) present a structural gravity model relying on monopolistic competition and CES consumer preferences, representing trade between countries as a function of the relative size of the two economies and the transportation frictions between them. Following this methodology, trade flows can be represented as:

$$X_{ijt} = K(Y_{it}Y_{jt})^{\alpha}\left(\frac{T_{ijt}}{P_iP_j}\right)^{1-\sigma} \tag{2.1}$$

where X_{ijt} is the total value of trade between country i and country j in year t, K is a constant, Y_{it} and Y_{jt} are the total outputs of country i and j respectively in year t, P_i and P_j are the price levels of country i and j respectively, and T_{ijt} represents transport costs between country i and country j. Costs, T_{ijt}, are modelled as 'iceberg' transport costs such that $T_{iit} = 1$ and $T_{ijt} \geq 1$.

The trade cost term T_{ijt} is usually modelled as a composite (often multiplicative) function of the various factors impacting the cost of trade including shipping costs, time, and uncertainty or insurance costs,

$$T_{ijt} = f(\tau_{ijt}, d_{ij}) \tag{2.2}$$

The distance term, d_{ij}, is used in order to proxy for many of the transportation and logistical costs as they tend to scale with the distance needing to be travelled. The τ_{ijt} effect often encompasses both fixed and time-varying border effects such as common institutional variables, trade agreements, contiguity, and any other impact on trade.

For simplicity, I assume that there are two countries, a small country, i, and the rest of the world, j. This collapses the function to one of total trade for country i with the potential for conflict in a neighbouring country.[7] With this assumption, transportation costs keep their standard form, but now represent a more generalised cost of trade with the rest of the world—more analogous to a market access measure—and therefore cannot be measured by distance. In this circumstance, time-invariant trade costs can be absorbed into country fixed effects. Similarly, the Y_{jt} and P_j terms from Equation 2.1 are no longer dependant on j and can be absorbed into year fixed effects.

Using this model structure, Equation 2.1 can be linearised with logs as follows

$$ln(X_{it}) = \alpha ln(Y_{it}Y_t) + ln(d_i) + ln(\tau_{it}) - ln(P_i)^{\sigma-1} - ln(P)^{\sigma-1} + \epsilon_{it} \tag{2.3}$$

This model inherits all of the features of standard structural gravity models including how to deal with the price indices P_i and P, the unobservable 'multi-lateral resistance' terms.[8] I include both country and year fixed effects in the estimation equations that follow to

[6]See Head and Mayer (2014); Feenstra (2004); Anderson and van Wincoop (2003).

[7]Adopting this structure for the model presents some difficulties by creating an aggregate 'trade partner' encompassing all the individual trade partners of a given country. Despite this, it benefits from simplifying the geographic elements used to identify the different effects that may be difficult to disentangle in a country-pair framework.

[8]See Anderson and van Wincoop (2003).

account for these unobservable aspects.[9] As mentioned, this combination of fixed effects absorbs the Y_t, d_i, and P terms.

I use the τ_{it} term to include variables for direct conflict within a country and adjacent conflict in a neighbouring country. Direct conflict is included as well in order to control for the direct impact of conflict on trade.[10] The supply- and demand-based effects are separately identified by exploiting the geographic features of countries as discussed in Section 2 and outlined below.

3.2 Direct and Adjacent Conflict

Using this model structure, the first estimation equation used in this analysis seeks to identify the generalised effect of adjacent conflict by using time variation in the exposure of one country to an adjacent conflict. Incorporating the fixed effects and conflict variables into Equation 2.3 gives:

$$ln(X_{it}) = \beta_1 ln(Y_{it}) + \beta_2 Conf_{it} + \beta_3 AdjConf_{it} + \rho_i + u_t + \epsilon_{it} \tag{2.4}$$

where X_{it} is the total trade (exports and imports)[11] for country i in year t, Y_{it} is the given country's real GDP in year t. Transport costs are covered by two sets of terms: the time-invariant country fixed effects (ρ_i) which encompasses the fixed aspects of a country's access to external markets and the two variables for conflict ($Conf_{it}$ and $AdjConf_{it}$). $Conf_{it}$ is an indicator variable for conflict within the observation country and $AdjConf_{it}$ is the main variable of interest: conflict in an adjacent country. Year fixed effects are given by u_t and ϵ_{it} is an error term. Standard errors in this equation and all following estimations are clustered at the region-year level.[12]

This equation includes the general impact of adjacent conflict on a country's trade, combining both the supply and demand effects. This gives some insight into how adjacent conflict impacts trade overall and provides a point of reference for later analyses of the separated effects.

The base model of adjacent conflict can be extended in two ways in line with the mechanisms discussed in Section 2 to better determine the supply effects of adjacent conflict. First, is country geography where there is likely to be a differentiated impact on landlocked countries as opposed to coastal countries due to a difference in ease of access to external markets. Second, is the geography of adjacent conflict with the potential impact of whether the conflict is taking place in an adjacent country that is shoreward or inland to the country of observation. Conflict in shoreward countries present a different risk to transportation routes than those that are inland to the country of observation due to their proximity to ports and transportation routes.

[9] As in Glick and Rose (2002), Anderson and van Wincoop (2004), and Baier and Bergstrand (2007).
[10] See Marano et al. (2013), Blomberg and Hess (2006), and Nitsch and Schumacher (2004) for more discussion of these direct effects.
[11] The effect of adjacent conflict on exports and imports separately is included in Appendix A.
[12] The regional breakdown of the world is given in Figure 2.5.

Country Geography A difference in impact of adjacent conflict on landlocked and coastal countries is represented by a modified version of Equation 2.4 which estimates separate coefficients for adjacent conflict according to whether a country is landlocked or coastal.

$$\begin{aligned} ln(X_{it}) = \beta_1 ln(Y_{it}) + \beta_2 Conf_{it} + \beta_3 AdjConf_{it} \times Coastal_i \\ + \beta_4 AdjConf_{it} \times Landl_i + \rho_i + u_t + \epsilon_{it} \end{aligned} \tag{2.5}$$

The variable $Landl_i$ is a dummy variable taking a value of 1 if the country is landlocked and 0 otherwise. $Coastal_i$ is a dummy variable that takes a value of 1 if the country is coastal and 0 otherwise.

The estimation in this form incorporates the differentiated effect of adjacent conflict interacting with the location of the country.[13] However, it does not impose that the conflict must happen along the coast and so identifies this general effect of adjacent conflict on countries where it is not disruptive (coastal countries) and countries where it is potentially disruptive (landlocked countries).

In the hypothetical example in Section 2, this equation captures the difference in the effect of the conflict in the first (adjacent) country for the second (landlocked) country and the third (coastal) country. In doing so, it allows for the separation of the supply effect from the demand effect as this difference between the impact of adjacent conflict on landlocked and coastal countries.

The difference between the supply and demand effects is given here by the difference in coefficients on $AdjConf_{it} \times Landl_i$ and $AdjConf_{it} \times Coastal_i$. This captures the difference caused by the supply effect in landlocked countries in addition to the more general demand effects across all countries.

Conflict Geography The second scenario involves the incorporation of two different adjacent conflict variables: one for conflict in an adjacent country that is shoreward; and one for conflict in an adjacent country that is inland.

$$\begin{aligned} ln(X_{it}) = {} & \beta_1 ln(Y_{it}) + \beta_2 Conf_{it} \\ & + \beta_3 AdjConf^S_{it} \times Coastal_i + \beta_4 AdjConf^S_{it} \times Landl_i \\ & + \beta_5 AdjConf^L_{it} \times Coastal_i + \beta_6 AdjConf^L_{it} \times Landl_i \\ & + \rho_i + u_t + \epsilon_{it} \end{aligned} \tag{2.6}$$

The superscripts attached to the $AdjConf_{it}$ variable indicate whether the conflict is taking place in an adjacent country that is shoreward (S) or inland (L).

This estimation captures the more detailed effect of the location of the adjacent conflict on trade. The difference between coastal and landlocked nations from Equation 2.5 remains

[13]The effect of direct conflict ($Conf_{it}$) interacted with the location of the country ($Landl_i$ or $Coastal_i$) is also performed. Direct conflict has a significantly more negative impact on trade in landlocked countries as compared to coastal countries, and landlocked countries remain more affected by adjacent conflict than coastal countries. These results are presented in Appendix A.

with the addition of whether the conflict itself takes place in a shoreward or inland country. This adds the last element discussed in the hypothetical case in Section 2, that the first (adjacent) country should be shoreward to the second (landlocked) country so that it has ports that the second (landlocked) country needs access to.

In this equation, the supply effect can then be separated from the demand effect by the difference between the coefficients on $AdjConf_{it}^{S} \times Landl_i$ and $AdjConf_{it}^{S} \times Coastal_i$. This represents the difference caused by disruption to trade routes and port access for the landlocked country.

3.3 Identification

The models presented in this chapter all incorporate country and year fixed effects, capturing constant unobserved heterogeneity across countries and years. The analysis exploits the country-time variation of adjacent conflict at the country level and so is unsuitable for the incorporation of fixed effects for time-varying unobserved heterogeneity across countries.

Two elements of omitted variation that are not easily captured by country and year fixed effects are country-specific and region-specific factors that could impact both trade flows and the propensity for conflict. Political institutions can influence the proliferation of conflict or the scale of a country's involvement in the world economy. Trade agreements can create trade relationships that may be more robust to the disruptions in trade flows presented by conflicts. Regional trends, such as growth or investment, can drive trade and suppress the potential for conflict. Accounting for these factors is discussed in the robustness checks presented in Section 5.2.

Related to this, is the potential for conflict to spill across borders causing countries not experiencing conflict to nevertheless be impacted by it. This creates the risk for an imperfect separation of countries that are affected by the conflict from those that are not. These spillover effects do not generally seem to manifest as additional conflicts, as the evidence in Section 4.3 appears to indicate; but this does not remove the possibility of an effect. The inclusion of the direct conflict variable in the estimation controls for this potential spillover in the country of observation. Wider spillover effects are more difficult to manage beyond region-specific effects. However, it is likely that if countries were to be negatively affected by conflict spillovers, it would bias the results downward leading to the underestimation of the effect of adjacent conflict on trade flows.

The question of reverse causality is of significance when dealing with issues of trade and conflict due to the potential for trade to reduce the incidence of conflict between countries.[14] Ideally, issues of reverse causation could be eliminated through the selection of conflicts that are truly exogenous to economic outcomes in the country. This would ensure that they are not derived from or ameliorated by trade flows or the possible wealth and growth generated by such flows. Unfortunately, this would likely exclude far too many

[14] See Oneal and Russet (1999a) and Hegre et al. (2010) for a discussion of this effect.

conflicts as the role of domestic economic outcomes in generating and sustaining conflict is well documented, particularly internal conflict.[15] Despite this, it is mostly a concern with the issue of direct conflict: conflict in the country of interest or various direct trade partners. The fact that this analysis deals primarily with adjacent conflict minimises this issue to some extent. Regardless, there remains the potential for some regional factors, such as wealth derived from trade flows, to reduce the incidence of conflict.

This then raises the issue of aggregation, since the structure of the data in this chapter means that the 'third-country' experiencing conflict is included in the aggregate trade flows of the observation country. This leaves open the possibility that the observed decrease in trade is actually driven by a decrease in bilateral trade between the country of observation and the country experiencing the conflict. The existence of an adjacent conflict effect in bilateral data (where the conflict country is truly a third-country) would dispel this possibility. A second solution is the removal of all direct trade with adjacent countries experiencing conflict when aggregating the trade partners. The results of both these solutions demonstrate that the effect is in fact one of adjacent conflict impacting trade with other countries and not a mis-categorisation of direct conflict.[16]

4 Data

4.1 Conflict

In order to measure conflict within countries (both direct and adjacent) I use the Major Episodes of Political Violence (MEPV) dataset (Marshall, 2019). This dataset includes measures for various types of conflict within and between all the countries listed in Table 2.15 since 1946. This chapter covers the period 1962 to 2012, well within the range covered in the MEPV data. These years are selected based on trade data availability and that by 1962 the majority of European colonial possessions were independent. Each type of conflict is coded on a scale from 0 to 10 with higher numbers indicating more intense conflict. There are four variables in the dataset that represent internal conflict: civil violence (CIVVIOL), civil war (CIVWAR), ethnic violence (ETHVIOL), and ethnic war (ETHWAR). Similarly, there are three variables that represent inter-state conflict: independence conflict (INTIND), international violence (INTVIOL), and international war (INTWAR). The dataset differentiates 'violence' and 'war' through the level of institutionalisation of the conflict—with the designation 'war' implying a higher institutionalisation and clearer strategic objectives.

The direct conflict variable used in this chapter is constructed from all four of the internal variables and the variable for independence conflicts.[17] The independence conflict measure was included for completeness in capturing conflict occurring within the territory

[15] Herbst (2000), Ballentine (2003), and Arnson and Zartman (2005) provide surveys of this tendency.
[16] Both of these results are presented in Appendix A.
[17] The results are not overly sensitive to adjustments in the composition of the conflict variables. Alternative constructions of the adjacent conflict variable are discussed in Section 5.2.

of a country. The two other international conflict measures were excluded in order to remove conflicts that are considered 'overseas' and do not involve violence within the territory of the country.

The various conflict types are summed and normalised to create the resulting conflict variable (dir_conf_i). The variable $conf_i$ used in the equations in Section 3.1 is a dummy variable indicating a value for the variable that is greater than 1:

$$conf_i = \begin{cases} 1 & \text{if } dir\_conf_i > 1; \\ 0 & \text{if } dir\_conf_i \leq 1. \end{cases}$$

A base threshold of 1 is used to remove small-scale demonstrations and internal conflicts that would be unlikely to have a large impact.[18]

To create the adjacent conflict variable, a procedure similar to that used in direct conflict is followed. The variable is therefore made up of civil violence (CIVVIOL), civil war (CIVWAR), ethnic violence (ETHVIOL), ethnic war (ETHWAR), and independence conflict (INTIND) which are combined in the same way as before to create an index of internal conflict (dir_conf_j) within a given neighbouring country j. The variable $adjconf_i$ used throughout is a dummy variable constructed using the following thresholds:

$$adjconf_i = \begin{cases} 1 & \text{if } dir\_conf_j > 1; \\ 0 & \text{if } dir\_conf_j \leq 1. \end{cases}$$

The same base threshold used in creating the variable for direct conflict is used here for consistency.

Neighbouring countries j for a given country i are derived from data on country adjacencies from the Correlates of War Direct Continuity dataset (Stinnett et al., 2002). The continuities in this dataset include direct land connections and a variety of sea distance proximities. For this chapter only direct land continuities are used in order to ensure that adjacent conflict captures the potential to disrupt trade routes and other supply factors—something that is less likely in countries separated by sea.

A different variation of the conflict variables mentioned above include conflict intensity measures. This formulation uses the conflict intensity scale included in the MEPV dataset as a measure of conflict intensity with higher values indicating more intense and/or destructive conflict.

Each individual category of conflict in the MEPV dataset uses a 10 point scale to score the level of conflict in that category in a given country during a given year. The intensity values themselves are based on several aspects of conflict beyond casualties and displacement since these can be heavily influenced by factors other than the type of conflict that occurs (notably poverty levels and the value placed on human life by command

[18] For example, protests associated with the Civil Rights Movement in the US and the 1968 Student Protests in France are coded as level 1 internal conflicts in the MEPV dataset. Level 2 is where conflict begins to be more sustained and disruptive, and less localised. Examples of level 2 conflicts include the Cuban Revolution and the Georgian Civil War of 1991-1993.

structures). These scales therefore include aspects such as relational goals, technology, and relative means of the groups involved in conflict. The MEPV dataset argues that this is a better means of determining the impact of conflict on society and this formulation is ideally suited to the questions addressed in this chapter since the measure cleanly captures state and infrastructure disruption levels.

The scale is comparable across types of conflict, time periods, and countries. Importantly, the conflict scale is also additive—two level-4 events are equivalent to a level 8 event. This allows multiple scores in different categories to be aggregated in a way that is meaningful and that this chapter takes advantage of in the construction of the conflict intensity variables. The categories and examples provided by the MEPV dataset have been reproduced in Table 2.14 for added clarity.

The intensity measure for adjacent conflict is created using the same base variables from the MEPV dataset. In this case, the intensity of conflict is summed across all conflict types (CIVVIOL, CIVWAR, ETHVIOL, ETHWAR, and INTIND) and all neighbouring countries j to create a total adjacent intensity measure for country i:

$$adjconf(int)_i = \sum_j \sum_x intensity_{xj}$$
$$\forall\, x \in \{\text{CIVVIOL, CIVWAR, ETHVIOL, ETHWAR, INTIND}\}$$

4.2 Other Variables

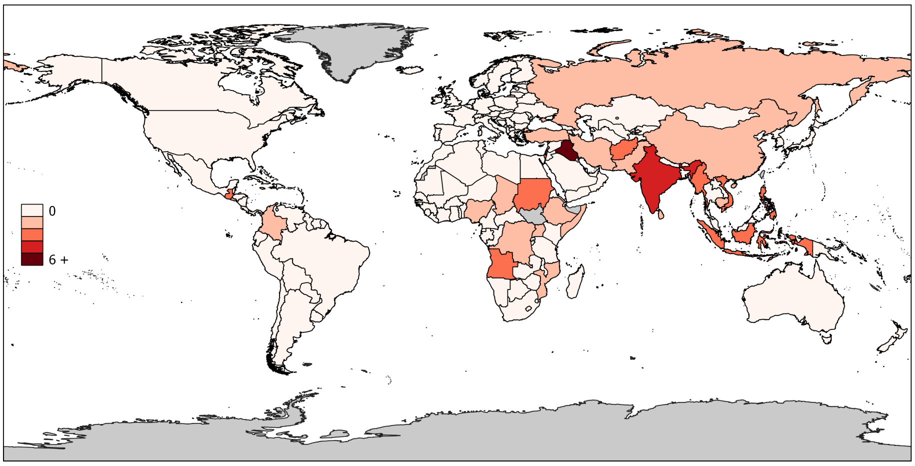
Trade data used in this analysis come from the UN COMTRADE database and include bilateral trade flows at the 4-digit SITC4 product code level. The years covered by the data are 1962 to 2012. This data was aggregated across products and partners to create a dataset of total trade in and out of a given country across all countries in the dataset.

The other variables used in the estimations include GDP and geographic aspects of the countries of interest. The landlocked or coastal nature of a country was determined by whether the country possessed coastline with an active port. As defined earlier, countries with a coast and port were coded as 'coastal' while countries without both these features were coded as 'landlocked'. Importantly, there are no countries in the dataset that have a coastline and no port. The GDP data for the various countries of interest (Y_{it}) come from World Bank figures. The list of countries used for this chapter is given in Table 2.15.

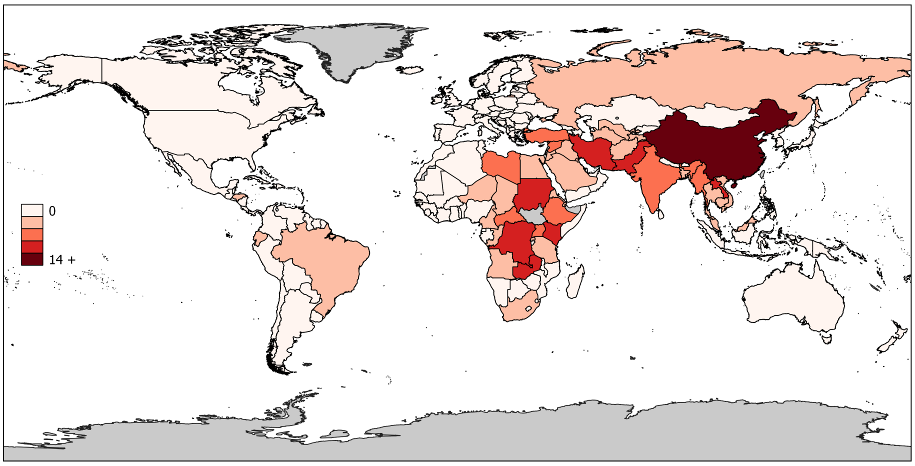
4.3 Descriptive Statistics

Figure 2.2 presents heat maps of conflict and adjacent conflict using intensity of conflict. These maps illustrate that particular continents tend to experience more conflict than others. Both Africa and Asia appear to experience the most and highest intensity conflict over the period of study. This is supported by Figure 2.3 which breaks down the incidence and intensity of conflict by continent. Africa and Asia tend to have both the highest

Figure 2.2. Intensity of Conflict



(a) Direct Conflict



(b) Adjacent Conflict

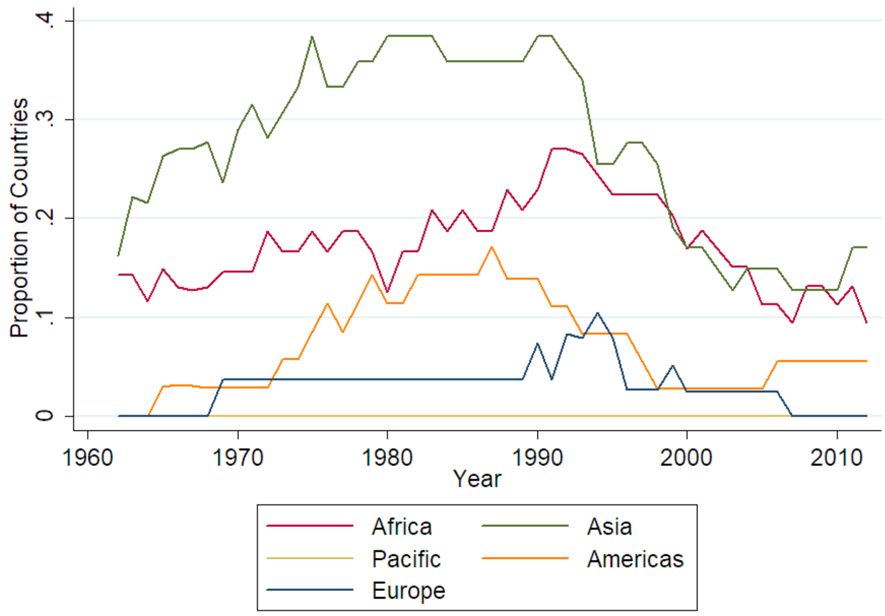
Note: Intensity of conflict (direct and adjacent) averaged across the period 1962-2012. Scales indicate the average conflict intensity, darker colours represent more intense conflict. (Source: Author's own calculations based on Marshall, 2019).

average incidence of conflict and intensity of conflict among the five major continental regions, with the Pacific experiencing the lowest.

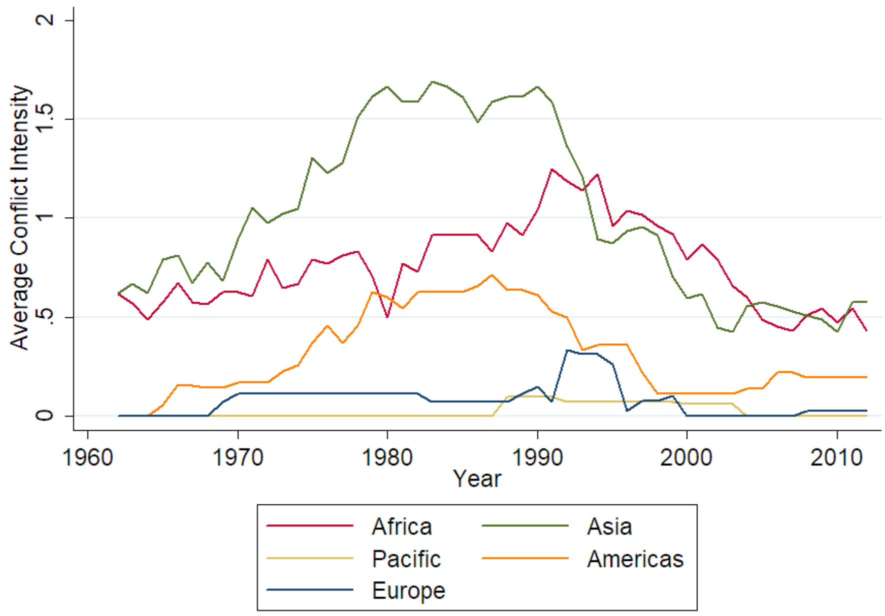
A year-by-year tracking of conflict incidence in the first panel of Figure 2.3 shows a general upward trend in all continents, peaking around the early 1990s and sharply falling off again. The continental breakdown of the chart shows that despite some continents consistently having a higher incidence of conflict than others, the trend is fairly uniform across them. The intensity of conflict also shown in the second panel follows a similar pattern, but peaks much earlier in most continents (except for Africa). While Asia pushes the intensity score upwards with a large peak during the 1980s, the general shape of global conflict intensity is consistent across continents with local peaks in different periods.

Adjacent conflict tends to be more common in the data than direct conflict; as Table

Figure 2.3. Incidence and Intensity of Conflict by Continent



(a) Countries in Conflict



(b) Intensity of Conflict

Note: Incidence and intensity of conflict by continent over the period 1962-2012. Lines in panel (a) represent the proportion of countries on each continent in conflict in the given year. Lines in panel (b) represent the average intensity of all conflicts in countries on each continent in the given year. (Source: Author's own calculations based on MEPV data from Marshall, 2019).

Table 2.1. Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Minimum	Maximum
ln(Trade)	8137	22.347	2.652	6.875	29.541
Conf	8137	0.132	0.339	0	1
AdjConf	8137	0.367	0.482	0	1
Conf (Intensity)	8137	1.462	0.010	0	10
AdjConf (Intensity)	8137	2.310	3.985	0	29

Note: Conf and AdjConf are dummy variables indicating the presence of direct conflict and adjacent conflict; Conf(Intensity) and AdjConf(Intensity) are variables that measure the intensity (given by the MEPV scale) of direct and adjacent conflict.

Table 2.2. Comparative Statistics

	Landlocked	Coastal	Difference
ln(Trade)	21.371 (0.064)	22.548 (0.032)	-1.177**
Conf	0.132 (0.009)	0.133 (0.004)	-0.001
AdjConf	0.555 (0.013)	0.329 (0.006)	+0.226**
Conf (Intensity)	0.528 (0.037)	0.525 (0.018)	+0.003
AdjConf (Intensity)	3.341 (0.114)	2.097 (0.047)	+1.244**

* Significant at 0.10; ** Significant at 0.05

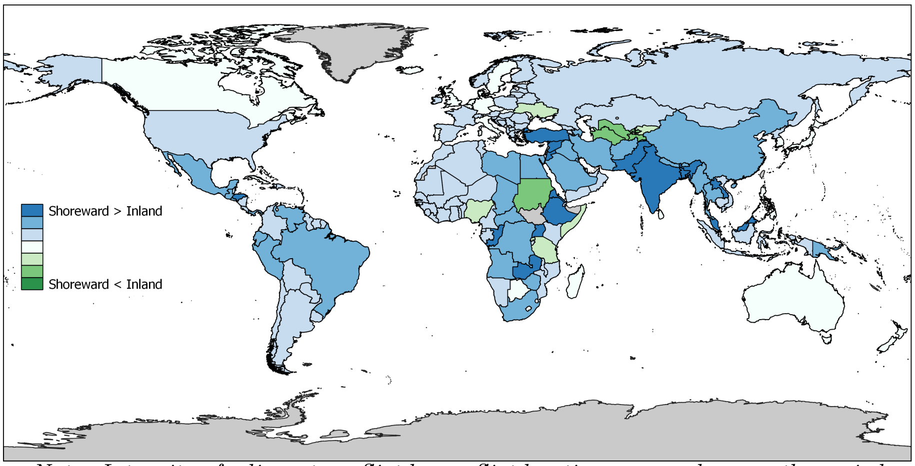
Note: This table presents the difference in means of ln(Trade), Conf, and AdjConf when a country is landlocked or coastal; standard errors are in parentheses; Conf and AdjConf are dummy variables, except where indicated as intensity measures.

2.1 shows, approximately 37% of the country-years covered in the data experience adjacent conflict compared to only 13% for direct conflict. This follows from the fact that a country in conflict in a given year will tend to have multiple neighbours that all experience the adjacent conflict in the same year. This difference in incidence of adjacent conflict compared to direct conflict combined with the differences in shading in Figure 2.2 between the map for direct conflict (the first panel) and the map for adjacent conflict (second panel), shows that there are many countries that are impacted by adjacent conflict that do not experience similar intensities of direct conflict themselves. This supports the idea that spillover effects are small by indicating that even if conflict does spill into neighbouring countries, it does so with significantly moderated intensity.

Landlocked countries tend to be more prone to adjacent conflict and more intense adjacent conflict than coastal countries, as indicated in Table 2.2. Importantly, this is not a general trend as landlocked countries do not seem to experience direct conflict at a higher rate or higher intensity when compared to coastal countries. This seems to indicate that landlocked countries tend to have multiple neighbours in conflict more frequently than coastal countries, producing a higher incidence but also a larger summed value of intensity than coastal countries.

The location of the conflicts themselves are displayed in Figure 2.4 which shows the propensity for adjacent conflict to occur in neighbouring shoreward nations or inland ones. Measured by intensity of conflict, the majority of adjacent conflict occurs in shoreward neighbouring nations. While this means that the vast majority of adjacent conflict in the dataset is likely to impact trade flows by interrupting trade routes, it also means that the sample size for inland adjacent conflict is relatively smaller making it more difficult to clearly identify effects. Additionally, this smaller sample size is likely to lead to larger standard errors, decreasing the potential statistical significance of results.

Figure 2.4. Intensity of Adjacent Conflict by Location



Note: Intensity of adjacent conflict by conflict location averaged across the period 1962-2012. Scale indicates the intensity and primary location of adjacent conflict for each country; blue represents primarily shoreward adjacent conflict, green represents primarily inland adjacent conflict. Darker colours represent more intense conflict. (Source: Author's own calculations based on Marshall, 2019).

5 Results

5.1 Main Results

The first column of Table 2.3 presents the results for the model given in Equation 2.4. The coefficients on both direct conflict and adjacent conflict are strongly negative and significant, with the effect of direct conflict being stronger than adjacent conflict. Direct conflict reduces trade by approximately 16.7% while adjacent conflict only reduces trade by approximately 6.8%. This is in line with the effects discussed in Section 3.1: conflict in an adjacent country seems to negatively impact trade, but not as severely as conflict directly in the country. Overall, these base results indicate that there is an adjacent conflict effect and it is both strong and significant.

The second column presents the results for Equation 2.5 which factors in the location of the country of observation (either landlocked or coastal). The coefficients on both interaction terms for adjacent conflict are negative and significant, with landlocked countries experiencing worse effects from conflict in adjacent countries. Conflict in an adjacent country reduces trade in a landlocked country by approximately 9.5%, almost twice as much as in a coastal country. While this difference in coefficients is not statistically significant, it provides some evidence that adjacent conflict seems to disrupt trade more for landlocked countries than coastal countries, lending support to the idea that the supply effect from disrupted trade routes and interrupted access to external markets discussed in Section 2 is both present and large.

Table 2.3. Principal Regression Results

	Incidence			Intensity	
	(1) Base Model	(2) Country Location	(3) Conflict Location	(4) Country Location	(5) Conflict Location
Conflict	-0.183** (0.029)	-0.184** (0.029)	-0.181** (0.029)	-0.178** (0.029)	-0.178** (0.029)
AdjConf	-0.070** (0.023)				
AdjConf x Coastal		-0.061** (0.023)		-0.007** (0.003)	
AdjConf x Landl		-0.100* (0.061)		-0.017* (0.009)	
AdjConf(S) x Coastal			-0.053** (0.022)		-0.008** (0.004)
AdjConf(S) x Landl			-0.124** (0.057)		-0.020** (0.010)
AdjConf(L) x Coastal			-0.125** (0.049)		-0.002 (0.006)
AdjConf(L) x Landl			0.034 (0.071)		-0.002 (0.016)
N	8137	8137	8137	8137	8137
R^2	0.964	0.964	0.964	0.964	0.964

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade; Conflict, Landl, and Coastal are dummy variables; AdjConf, AdjConf(S), and AdjConf(L) measure the presence (first three columns) or intensity (last two columns) of adjacent conflict, shoreward adjacent conflict, and inland adjacent conflict respectively. The log value of country GDP is included as an independent variable.

The third column displays the result for the model corresponding to Equation 2.6, combining the location of the conflict (shoreward or inland) as well as the location of the country of observation (coastal or landlocked). The interaction terms demonstrate clearly that adjacent conflict in a shoreward country is significantly worse for a landlocked country than a coastal country. Adjacent conflict in shoreward countries appears to reduce trade by 11.6% for landlocked countries, over twice the impact on coastal countries. Interestingly, the coefficient for inland conflict for coastal countries is also strongly negative, while it is indistinguishable from zero for landlocked countries. This appears to indicate that inland conflict is very damaging for coastal countries. This may be influenced by the relative lack of data for landlocked countries (and inland conflict) and so might be driven by outliers since the magnitude of the effect shrinks considerably and is insignificant with the use of a continuous measure of conflict in the fifth column.

These results help to illustrate the difference between the supply effect and the demand effect for adjacent conflict, demonstrating the larger impact of the supply effect on landlocked countries which are more sensitive to disruptions in trade routes and restricted access to external markets through other countries' ports. The difference in estimated magnitude between coefficients on adjacent conflict for landlocked and coastal countries in column 2 helps to disentangle the supply effect from the demand effect, and provides intuition for the impact of disruption to infrastructure and market access on trade flows. While coastal nations have their own ports to access external markets, landlocked countries do not and the larger impact of adjacent conflict demonstrates the difficulty that these countries have in accessing ports in other countries with conflict along the routes. This is supported by the results in column 3 where the location of the conflict is separated out. The impact of shoreward adjacent conflict specifically on the trade of landlocked countries is more than two times higher than for coastal countries. This emphasises the impact of adjacent conflict on access to external markets.

The final two columns of Table 2.3 present the results for the models given in Equations 2.5 and 2.6 using the conflict intensity measure discussed in Section 4.1. These results reinforce the conclusions about the impact of adjacent conflict on countries and trade.

The fourth column presents results for a continuous measure of adjacent conflict using the specification in Equation 2.5, analogous to the second column. This result also provides evidence that landlocked countries appear to experience increasingly worse effects of adjacent conflict than coastal countries, with trade reducing by 1.7% for each unit increase in conflict intensity in neighbouring countries (as opposed to a decrease of 0.7% for coastal countries). Despite this difference not being statistically significant, this provides some insight that landlocked countries appear to be more sensitive to conflict intensity—suffering a reduction of trade more than twice as large as coastal countries for each unit increase in conflict intensity.

The final column presents the results for the location of conflict model (as in column 3) for the conflict intensity measure. These results also show that landlocked countries are more sensitive to shoreward conflict than coastal countries, experiencing a 2% decline in trade for each increase in conflict intensity. By contrast, coastal countries experience only a 0.8% drop in trade, less than half the effect of landlocked countries. The results for inland conflict are less instructive as neither coefficient is significant and they are not different in magnitude.

These results for the intensity measure add a subtlety to the effect revealed by the dummy variables. The presence of adjacent conflict, and shoreward adjacent conflict in particular, appears to be worse for landlocked countries than coastal countries; further, this negative effect increases with the intensity of the adjacent conflict. The larger magnitude of the coefficients on landlocked countries as opposed to coastal countries (in column 4 and for shoreward conflict in column 5) seem to indicate that the negative supply effects from adjacent conflict grow at a faster rate than the demand effects. This makes sense from an intuitive standpoint: the demand effects are largely determined by the presence or perception of conflict (boycotts or substitutions) and are possibly less sensitive to small changes in intensity, while the supply effects are grounded in infrastructural impacts which scale with conflict intensity.

5.2 Robustness

I conduct a number of robustness checks on the results presented above. The robustness results deal with two primary aspects of the analysis: the first is model identification issues such as omitted variation, while the second is the construction of the conflict variable. A third element includes spatially correlated standard errors.

Endogeneity Concerns There are many country-specific and regional effects that can ultimately impact both trade and conflict. This is even more the case in this analysis, which looks at conflict in neighbouring countries. Elements such as political stability, trading blocs, regional growth trends, and various other political and economic variables can have wider impacts on both conflict and trade flows within a given region. This effect is also linked to issues of reverse causation in that regional trade flows and trade-driven growth can reduce the incidence of regional conflict.

I account for two examples of these country-specific omitted effects with the inclusion of measures for political stability and trade agreements. Political institutions within a country can have an effect on how much it trades with other countries and the resilience of these connections to stress and shocks. In order to account for this potential effect, I include the Polity IV (Marshall et al., 2018) score as a control for the political institutions of the country of observation. This score measures the openness and inclusivity of the political institutions in a country, assigning a score between -10 and $+10$ with higher numbers indicating a more democratic and inclusive institutional structure.

In addition to this, I include a control for the number of trade agreements that the country of observation is a party to using data from the WTO database. Trade agreements

Table 2.4. Political Stability and Trade Agreement Controls

	(1) Country Location	(2) Conflict Location	(3) Country Location	(4) Conflict Location
Conflict	-0.171**	-0.169**	-0.169**	-0.167**
	(0.027)	(0.027)	(0.027)	(0.027)
AdjConf	-0.084**		-0.080**	
x Coastal	(0.021)		(0.020)	
AdjConf	-0.123**		-0.131**	
x Landl	(0.058)		(0.060)	
AdjConf(S)		-0.090**		-0.086**
x Coastal		(0.020)		(0.020)
AdjConf(S)		-0.154**		-0.161**
x Landl		(0.055)		(0.056)
AdjConf(L)		-0.071		-0.068
x Coastal		(0.043)		(0.044)
AdjConf(L)		0.034		0.051
x Landl		(0.074)		(0.072)
RTAs			0.005**	0.005**
			(0.001)	(0.001)
Polity Score			0.014**	0.014**
			(0.002)	(0.002)
N	6652	6652	6652	6652
R^2	0.967	0.967	0.968	0.968

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade; Conflict, Landl, and Coastal are dummy variables; AdjConf, AdjConf(S), and AdjConf(L) are dummy variables that measure the presence of adjacent conflict, shoreward adjacent conflict, and inland adjacent conflict respectively. RTAs is the number of trade agreements a country is a member of; Polity Score is the country's Polity IV value indicating strength of domestic institutions. The log value of country GDP is included as an independent variable.

Table 2.5. Regional Time Trends

	(1) Base Model	(2) Country Location	(3) Conflict Location
Conflict	-0.170**	-0.170**	-0.170**
	(0.031)	(0.031)	(0.031)
AdjConf	-0.043**		
	(0.021)		
AdjConf x Coastal		-0.041**	
		(0.020)	
AdjConf x Landl		-0.053	
		(0.059)	
AdjConf(S) x Coastal			-0.030
			(0.019)
AdjConf(S) x Landl			-0.076
			(0.054)
AdjConf(L) x Coastal			-0.055
			(0.047)
AdjConf(L) x Landl			0.034
			(0.065)
N	8137	8137	8137
R^2	0.968	0.968	0.968

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade; Conflict, Landl, and Coastal are dummy variables; AdjConf, AdjConf(S), and AdjConf(L) are dummy variables indicating adjacent conflict, shoreward adjacent conflict, and inland adjacent conflict. The log value of country GDP is included as an independent variable. Regional time trends are included in all specifications.

can create robust trade links to other countries that could counteract the adverse effects of adjacent conflict.

The results for the inclusion of these controls are presented in Table 2.4. The first two columns present the baseline estimations with the newly reduced sample size, the second two columns present the results for the inclusion of the new variables. The inclusion of these controls appears to slightly strengthen the findings on adjacent conflict, contributing stronger evidence to the suggestion of a supply effect of adjacent conflict. Column 3 demonstrates that with these controls included, the magnitudes attached to landlocked countries become more negative relative to coastal countries. Further, the effect of shoreward conflict on landlocked countries in the fourth column is stronger relative to the impact on coastal countries than in the base analysis.

Despite the inclusion of these examples, this does not capture the full extent of potential regional effects that can be associated with both trade and conflict. In order to account for these more generally, I include a regional time trend within the estimation. These results are presented in Table 2.5. The incorporation of the regional time trend inherently removes some of the variation in the geographic variables and so the significance of the results are reduced, this is particularly true for landlocked countries since there are already few cases within the data. This makes it difficult to separate the broader regional effects from the more specific neighbour effects looked at in this chapter. Despite this, the magnitudes remain suggestive of the supply effect discussed previously.

Conflict Variables The second set of robustness checks deal with adjustments to the measurement of adjacent conflict in the data. This includes a measure of adjacent conflict that accounts for the number of neighbouring countries and variations in the variables used to construct the overall conflict measure.

While the main analysis uses geographic elements of conflict and countries themselves in an effort to disentangle the competing effects of adjacent conflict on trade, there is a secondary element of geography that can be added to the measure of conflict by moderating the intensity of adjacent conflict by the number of neighbours the country has. To do this, I create a new measure of adjacent conflict, building on the previously constructed measure of adjacent conflict intensity:

$$adjconf(int)_i^n = \frac{\sum_j \sum_x intensity_{xj}}{J}$$
$$\forall\, x \in \{\text{CIVVIOL, CIVWAR, ETHVIOL, ETHWAR, INTIND}\}$$

Where conflict intensity is summed across all conflict types, x, and all neighbouring countries, j, and divided by the number of neighbouring countries, J. The n superscript denotes the variable as moderated by the number of neighbours. This measure of adjacent conflict intensity moderates the conflict measure by the number of neighbours, giving an indication of not only the intensity of adjacent conflict, but also the incidence of it in existing neighbours. This captures a different geographic element of adjacent conflict by accounting for the number of potential conflicts in neighbouring countries.

Table 2.6. Adjacent Conflict and Neighbouring Countries

	(1) Base Model	(2) Country Location	(3) Conflict Location
Conflict	-0.181**	-0.260**	-0.252**
	(0.029)	(0.030)	(0.031)
AdjConf	-0.125**		
	(0.044)		
AdjConf x Coastal		-0.020	
		(0.046)	
AdjConf x Landl		-0.392**	
		(0.162)	
AdjConf(S) x Coastal			0.051
			(0.044)
AdjConf(S) x Landl			-0.399**
			(0.137)
AdjConf(L) x Coastal			-1.141**
			(0.213)
AdjConf(L) x Landl			-0.683*
			(0.388)
N	8137	8137	8137
R^2	0.964	0.947	0.947

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade; Conflict, Landl, and Coastal are dummy variables; AdjConf, AdjConf(S), and AdjConf(L) measure the intensity of conflict in neighbouring countries divided by the number of neighbours a country has. The log value of country GDP is included as an independent variable.

The results for this measure of adjacent conflict are presented in Table 2.6. The results are largely in line with the baseline analysis, though the coefficients on adjacent conflict for coastal countries are no longer significant, strengthening the results for landlocked countries in comparison and adding further evidence for the existence of a supply effect in the impact of adjacent conflict on countries.

Table 2.7. Alternate Conflict Variables

	(1) Base Conflict	(2) Civil Conflict	(3) Ethnic Conflict	(4) Ethnic Violence	(5) Ethnic War
Conflict	-0.184**	-0.185**	-0.186**	-0.179**	-0.185**
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
AdjConf x Coastal	-0.061**	-0.063**	-0.146**	-0.118**	0.009
	(0.023)	(0.023)	(0.027)	(0.030)	(0.027)
AdjConf x Landl	-0.100*	-0.096	-0.167**	-0.226**	-0.018
	(0.061)	(0.061)	(0.062)	(0.061)	(0.068)
N	8137	8137	8137	8137	8137
R^2	0.964	0.964	0.964	0.964	0.964

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade; Conflict, Landl, and Coastal are dummy variables; AdjConf is a dummy variable that measures the presence of conflict (as defined in each column) in neighbouring countries. The log value of country GDP is included as an independent variable.

The second check adjusts the variables used to calculate the measure of adjacent conflict. Presented in Table 2.7 are four additional measures of the conflict variable, the first column presents the measure used in the principal analysis for reference. The first measure, Civil Conflict, includes the measures CIVVIOL and CIVWAR. The second measure combines only the variables for ethnic-based conflict, ETHVIOL and ETHWAR. The final two columns are the variables ETHVIOL and ETHWAR by themselves.

All four measures produce results with magnitudes that are in line with those found in the main analysis. While two of the measures produce results that are not significant, the measures for Ethnic Conflict and Ethnic Violence are significant. This seems to indicate that ethnic-based violence, which is more disorganised than war and potentially problematic for trade routes and logistics, is more impactful on trade than more structured types of conflict (ethnic war for instance) and other types of civil conflict.

Inference Whereas the baseline results included standard errors clustered at the region-year level, the results in Table 2.8 refine this to include spatially correlated standard errors, following Conley (2008). The implementation of this level of clustering produces significant results that remain suggestive of the supply effect with landlocked countries continuing to experience more negative effects than coastal countries, particularly in the case of shoreward adjacent conflict.

Table 2.8. Spatial Standard Errors

	(1) Base Model	(2) Country Location	(3) Conflict Location
Conflict	-0.183**	-0.184**	-0.181**
	(0.037)	(0.037)	(0.037)
AdjConf	-0.070**		
	(0.026)		
AdjConf x Coastal		-0.061**	
		(0.028)	
AdjConf x Landl		-0.100*	
		(0.061)	
AdjConf(S) x Coastal			-0.053**
			(0.027)
AdjConf(S) x Landl			-0.124**
			(0.059)
AdjConf(L) x Coastal			-0.125**
			(0.057)
AdjConf(L) x Landl			0.034
			(0.076)
N	8137	8137	8137
R^2	0.405	0.405	0.406

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade; Conflict, Landl, and Coastal are dummy variables; AdjConf, AdjConf(S), and AdjConf(L) are dummy variables indicating adjacent conflict, shoreward adjacent conflict, and inland adjacent conflict. The log value of country GDP is included as an independent variable. Errors are clustered spatially following Conley (2008).

6 Conclusion

This chapter provides evidence for the impact of adjacent conflict on trade flows through disruption to trade routes and access to external markets. Using data on conflict to generate variables for both direct and adjacent conflict within countries, I exploit the geographic features of both countries and conflict to illustrate the different impacts of adjacent conflict on landlocked and coastal countries. The results indicate the presence of this supply-driven effect, particularly for landlocked countries experiencing shoreward adjacent conflict.

While the scarcity of some data prevents a fuller understanding of what is driving the effects, the results are indicative of the negative effects of adjacent conflict through the supply channel outlined in Section 2. The focus of this chapter on trade flows provides evidence for an important aspect of the economic effects of adjacent conflict, but expanding beyond this into the implications of these trade effects for the wider economy could create a broader understanding of the effects of adjacent conflict. Similarly, a fine-grained analysis of the strategies used by firms and countries to ameliorate the effects of adjacent conflict—either through redirection of goods or different choices of transportation method—could supplement the work in the other direction.

Likely the only convincing way to definitively solve the problem of reverse causation is to select a set of truly exogenous conflicts—those conflicts that have no connection to economic sources. This is fundamentally difficult to extract; since wars are costly, they will frequently be fought for economic gain. However, if this could be done it would go a long way to solving the endogeneity problems inherent in analysing the impact of conflict on economic outcomes.

The insights into the effect of adjacent conflict in this chapter are very relevant to research in trade and peace studies, particularly in terms of the interaction between regional conflict and trade flows. The outline of the effect of conflict via a supply channel presented in this chapter opens up new directions in analysis of the various impacts of adjacent conflict. Going beyond traditional understandings of direct spillover effects into the more indirect effects such as the demand channel outlined in Qureshi (2013) and the supply channel discussed here allows for a more complete picture of the impact of regional conflicts on the economies of non-participants. This work supports the results of other studies in the effects of conflict on trade and contributes to the growing consensus on the negative impacts of adjacent conflict.

A Appendix: Supplementary Regressions

This appendix contains three supplementary estimations for the main analysis to answer two questions. The first differentiates between direct conflict in coastal and landlocked countries in a manner analogous to that used for adjacent conflict in the main analysis to determine if the supply effect identified in the main analysis also plays a role in the effect of direct conflict on trade. The second is the question of what is driving the decrease in trade due to adjacent conflict in the aggregated trade data. Two estimations are included to exclude the possibility that this decrease is due to lower trade with the country experiencing the conflict.

A.1 Interacting Direct Conflict

The main analysis in this chapter addresses the effect of geography on adjacent conflict and trade, but the effects of direct conflict can also be differentiated according to the location of the country. It is possible that the trade of landlocked countries could be more affected by domestic conflict than coastal countries for similar reasons to that of adjacent conflict.

The increased logistics necessary to export from and import to a landlocked country makes trade routes more fragile to domestic disruption. While some firms in coastal countries may still be able to access ports due to proximity or the specific locality of the conflict, landlocked countries must trade through other nations. Domestic conflict can create both direct and indirect barriers to this activity: the conflict can directly disrupt domestic infrastructure, making it more difficult to move products within the country; neighbouring countries may also close border crossings in an effort to contain the violence, interrupting the access of landlocked countries to their ports.

The results in this subsection estimate the following equation:

$$\begin{aligned} ln(X_{it}) = \beta_1 ln(Y_{it}) + \beta_2 Conf_{it} \times Coastal_i + \beta_3 Conf_{it} \times Landl_i \\ + \beta_4 AdjConf_{it} \times Coastal_i + \beta_5 AdjConf_{it} \times Landl_i + \rho_i + u_t + \epsilon_{it} \end{aligned} \tag{2.7}$$

The variable $Landl_i$ is a dummy variable taking a value of 1 if the country is landlocked and 0 otherwise. $Coastal_i$ is a dummy variable that takes a value of 1 if the country is coastal and 0 otherwise. $Conf_{it}$ and $AdjConf_{it}$ are dummy variables indicating the presence of direct and adjacent conflict respectively.

Table 2.9 presents the results for including the geographic effects of direct conflict. The first column reproduces the results of the main analysis with an undifferentiated direct conflict measure. The second column splits this effect for landlocked and coastal countries. The results indicate that landlocked countries do experience significantly worse effects from domestic conflict than coastal countries, reducing trade by almost 33% compared to 12% for coastal countries.

Table 2.9. Direct Conflict Interaction Results

	(1)	(2)
Conflict	-0.184**	
	(0.029)	
Conflict x Coastal		-0.124**
		(0.028)
Conflict x Landl		-0.396**
		(0.083)
AdjConf x Coastal	-0.061**	-0.066**
	(0.023)	(0.023)
AdjConf x Landl	-0.100*	-0.106*
	(0.061)	(0.061)
N	8137	8137
R^2	0.964	0.964

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade; Conflict and AdjConf are dummy variables that indicate the presence of direct conflict and adjacent conflict; Landl and Coastal are dummy variables for whether the country is land-locked or coastal. The log value of country GDP is included as an independent variable.

This seems to indicate that the supply effect outlined for adjacent conflict in the main analysis also has a role to play in direct conflict. Further the effect appears to be very pronounced for direct conflict, nearly tripling the effect of domestic conflict on trade.

A.2 Effect of Adjacent Conflict on Aggregate Trade

The aggregation of trade data in this chapter leaves open the possibility that the decrease in trade due to adjacent conflict is caused by a fall in trade with the adjacent country itself. As discussed in the main analysis, there are two ways to address this problem. The first is the use of country-pair data where the adjacent country experiencing conflict is outside the trade relationship and therefore cannot influence trade flows directly. The second is the exclusion of neighbouring countries from the aggregated trade calculations. Both these methods are given below.

Country-Pair Analysis Using bilateral trade data where the adjacent conflict is truly a third country can help to address the issue of the cause of the decrease in trade. With bilateral data, there is no possibility that any decrease in trade due to adjacent conflict is the result of lost trade with the country in conflict.

The following estimates a traditional bilateral gravity model with an estimation equation of the following form:

$$\begin{aligned} ln(X_{it}) &= \beta_1 ln(Y_{it}Y_{jt}) + \beta_2 ln(d_{ij}) + \beta_3 Landl_{ij} \\ &+ \beta_4 Conf_{ijt} + \beta_5 AdjConf_{ijt} + \rho_i + \rho_j + u_t + \epsilon_{it} \end{aligned} \tag{2.8}$$

Where Y_{it} and Y_{jt} are the GDP of countries i and j, d_{ij} is the distance between them, and $Landl_{ij}$ indicates if one of the countries in the pair is landlocked. $Conf_{it}$ and $AdjConf_{it}$ are dummy variables for if one of the countries in the pair is experiencing direct conflict or adjacent conflict (in a country outside the pair).

The data used for this estimation are the un-aggregated bilateral data in the main analysis. The trade data are therefore at the 4-digit SITC4 product level. All other data sources are the same.

The results in Table 2.10 show that adjacent conflict does impact bilateral trade flows, reducing trade by approximately 8%. This indicates that the effect of adjacent conflict in general is not one of reduced trade with the country experiencing conflict. The conflict in a third country significantly impacts trade between a neighbouring country and its trade partners.

Table 2.10. Country-Pair Results

	(1)
ln(GDP)	0.459**
	(0.005)
ln(Dist)	-1.699**
	(0.004)
Landl	-0.644**
	(0.020)
Conflict	-0.144**
	(0.009)
AdjConf	-0.088**
	(0.007)
N	757734
R^2	0.688

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade; ln(GDP) is log value of product of country GDPs; ln(Dist) is log distance between countries; Conflict and AdjConf are dummy variables indicating direct conflict and adjacent conflict for at least one of the countries in the country-pair.

Excluding Neighbouring Trade The second method of addressing this issue is the exclusion of adjacent countries from the aggregate trade figures. By removing these countries, there is no potential for a reduction in trade with the adjacent conflict country impacting aggregate trade and there can be certainty that the effect identified is one of adjacent conflict impacting trade flows with non-conflict countries.

The results presented in Table 2.11 reproduce the results of the main analysis, but with a modified calculation of aggregate trade flows that excludes neighbouring countries. The first column shows that adjacent conflict still has a similar effect on the modified calculation of aggregate trade. The second column shows the results for adjacent conflict differentiated across landlocked and coastal countries. The result for coastal countries is significant and similar to that in the main analysis, but the result for landlocked countries is much smaller and insignificant. This casts some doubt on the effect of adjacent conflict for landlocked countries and the size of the supply effect.

The results for the third column are similar to those of the second, the results for coastal countries are significant and similar in magnitude to those in the main analysis but the results for landlocked countries are insignificant.

Ultimately, excluding neighbouring countries from the calculations of aggregate trade means that the estimates for landlocked countries lose statistical significance. This could be related to the fact that landlocked countries primarily trade with neighbouring countries and so much of the impact is felt through those trade links. This could be particularly relevant in the positive coefficient for landlocked countries with inland adjacent conflict. This value could represent the redirection of trade to other countries, potentially increasing overall trade flows, in an effort to bypass the conflict zone.

A.3 Splitting Exports and Imports

The aggregation of trade in the main analysis does not explicitly distinguish between exports and imports. It is possible that exports and imports are differentially affected by adjacent conflict, with one direction experiencing more acute impacts of conflict in neighbouring countries than the other. Tables 2.12 and 2.13 replicate the specification from the main results but replace the aggregated trade variable with exports and imports respectively.

The results for exports given in Table 2.12 are stronger than those of the main analysis, with the coefficients increasing in magnitude and significance. The first three columns present the results for the incidence measure of adjacent conflict while the final two columns show the results for adjacent conflict intensity. The increase in magnitude and significance is common to both measures of adjacent conflict, demonstrating that exports are strongly affected by conflict in neighbouring countries.

In contrast, the results for imports in Table 2.13 decrease in both magnitude and significance from the main results. Again, the first three columns present the results for the incidence measure of adjacent conflict and the final two columns display the intensity

Table 2.11. Excluding Neighbouring Trade

	(1) Base Model	(2) Country Location	(3) Conflict Location
Conflict	-0.179** (0.029)	-0.178** (0.029)	-0.177** (0.029)
AdjConf	-0.065** (0.022)		
AdjConf x Coastal		-0.072** (0.023)	
AdjConf x Landl		-0.037 (0.055)	
AdjConf(S) x Coastal			-0.062** (0.021)
AdjConf(S) x Landl			-0.077 (0.053)
AdjConf(L) x Coastal			-0.091* (0.048)
AdjConf(L) x Landl			0.081 (0.071)
N	8137	8137	8137
R^2	0.961	0.961	0.961

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Trade), log value of total trade excluding all trade with neighbouring countries; Conflict, Landl, and Coastal are dummy variables; AdjConf, AdjConf(S), and AdjConf(L) are dummy variables that measure the presence of adjacent conflict, shoreward adjacent conflict, and inland adjacent conflict respectively. The log value of country GDP is included as an independent variable.

Table 2.12. Export Trade

	Incidence			Intensity	
	(1) Base Model	(2) Country Location	(3) Conflict Location	(4) Country Location	(5) Conflict Location
Conflict	-0.271**	-0.271**	-0.266**	-0.257**	-0.258**
	(0.033)	(0.033)	(0.032)	(0.033)	(0.032)
AdjConf	-0.134**				
	(0.027)				
AdjConf x Coastal		-0.135**		-0.019**	
		(0.028)		(0.003)	
AdjConf x Landl		-0.130**		-0.042**	
		(0.064)		(0.009)	
AdjConf(S) x Coastal			-0.116**		-0.019**
			(0.026)		(0.004)
AdjConf(S) x Landl			-0.189**		-0.054**
			(0.063)		(0.010)
AdjConf(L) x Coastal			-0.254**		-0.019**
			(0.057)		(0.008)
AdjConf(L) x Landl			0.141*		0.009
			(0.083)		(0.017)
N	8125	8125	8125	8125	8125
R^2	0.953	0.953	0.953	0.953	0.954

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Exports), log value of total exports; Conflict, Landl, and Coastal are dummy variables; AdjConf, AdjConf(S), and AdjConf(L) measure the presence (first three columns) or intensity (last two columns) of adjacent conflict, shoreward adjacent conflict, and inland adjacent conflict respectively. The log value of country GDP is included as an independent variable.

measure. The results lose much of their significance across both measures, indicating that imports are not overly affected by conflict in neighbouring countries. Despite this lack of significance, the signs of the coefficients are still largely negative suggesting the potential for some negative effects from adjacent conflict.

The combination of these results for exports and imports show that exports are much more impacted by conflict in adjacent countries than imports. This suggests a differentiated impact of the supply effect between the two directions of trade. This may be explained by the difference in ease of substitution between exports and imports. It is relatively easy and straightforward for partner countries to substitute away from the exports of a country experiencing adjacent conflict—sourcing their imports from alternative exporters. In contrast, there are no conflict-free alternative routes for imports into a country experiencing adjacent conflict and so import trade continues to flow, though potentially with increased friction as the signs of the coefficients seem to indicate.

Table 2.13. Import Trade

	Incidence			Intensity	
	(1) Base Model	(2) Country Location	(3) Conflict Location	(4) Country Location	(5) Conflict Location
Conflict	-0.162**	-0.163**	-0.160**	-0.161**	-0.160**
	(0.031)	(0.031)	(0.032)	(0.031)	(0.031)
AdjConf	-0.049**				
	(0.025)				
AdjConf x Coastal		-0.040*		-0.002	
		(0.023)		(0.003)	
AdjConf x Landl		-0.080		-0.004	
		(0.069)		(0.009)	
AdjConf(S) x Coastal			-0.035		-0.003
			(0.022)		(0.004)
AdjConf(S) x Landl			-0.091		-0.003
			(0.064)		(0.010)
AdjConf(L) x Coastal			-0.090*		0.003
			(0.050)		(0.006)
AdjConf(L) x Landl			-0.045		-0.008
			(0.074)		(0.017)
N	8135	8135	8135	8135	8135
R^2	0.960	0.960	0.960	0.960	0.960

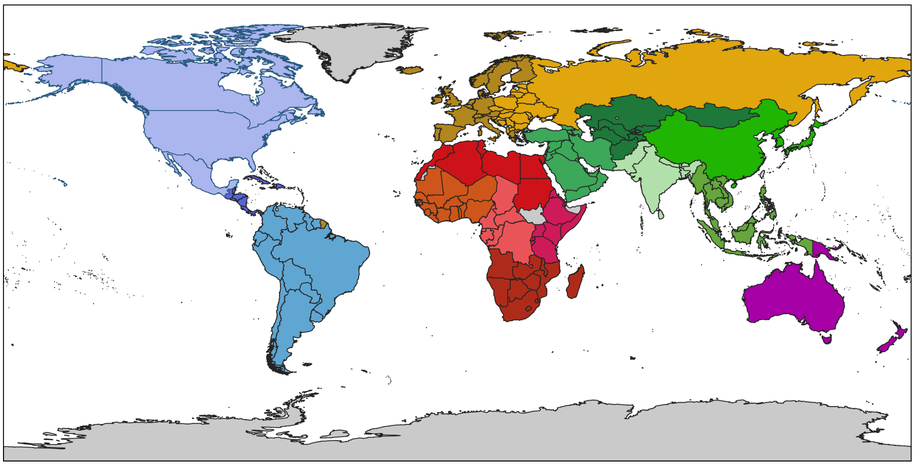
Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is ln(Imports), log value of total imports; Conflict, Landl, and Coastal are dummy variables; AdjConf, AdjConf(S), and AdjConf(L) measure the presence (first three columns) or intensity (last two columns) of adjacent conflict, shoreward adjacent conflict, and inland adjacent conflict respectively. The log value of country GDP is included as an independent variable.

B Appendix: Supplementary Figures and Tables

Figure 2.5. Regional Country Groupings



Note: Regional country groupings are indicated by colour.

Table 2.14. Conflict Intensity Definitions

Intensity	Definition	Example
1	Sporadic or Expressive Political Violence	United States, 1965-68
2	Limited Political Violence	Georgia, 1991-93
3	Serious Political Violence	Chile, 1974-76
4	Serious Warfare	Liberia, 1990-97
5	Substantial and Prolonged Warfare	Guatemala, 1966-96
6	Extensive Warfare	Bosnia, 1992-95
7	Pervasive Warfare	Vietnam, 1958-75
8	Technological Warfare	France, 1914-18
9	Total Warfare	Germany, 1941-45
10	Extermination and Annihilation	Japan, 1945

Source: MEPV Dataset (Marshall, 2019)

Table 2.15. List of Countries

Afghanistan	Denmark	Laos	Saint Kitts and Nevis
Albania	Djibouti	Latvia	Saint Lucia
Algeria	Dominica	Lebanon	Saint Vincent and the Grenadines
American Samoa	Dominican Republic	Lesotho	Samoa
Andorra	Ecuador	Liberia	San Marino
Angola	Egypt	Libya	Sao Tome and Principe
Antigua and Barbuda	El Salvador	Lithuania	Saudi Arabia
Argentina	Equatorial Guinea	Luxembourg	Senegal
Armenia	Eritrea	Madagascar	Serbia
Aruba	Estonia	Malawi	Sewden
Australia	Eswatini	Malaysia	Seychelles
Austria	Ethiopia	Maldives	Sierra Leone
Azerbaijan	Fiji	Mali	Singapore
Bahamas	Finland	Malta	Slovakia
Bahrain	France	Marshall Islands	Slovenia
Bangladesh	Gabon	Mauritania	Solomon Islands
Barbados	Gambia	Mauritius	Somalia
Belarus	Georgia	Mexico	South Africa
Belgium	Germany	Micronesia	Spain
Belize	Ghana	Moldova	Sri Lanka
Benin	Greece	Mongolia	Sudan
Bhutan	Grenada	Morocco	Surinam
Bolivia	Guam	Mozambique	Switzerland
Bosnia and Herzegovina	Guatemala	Myanmar	Syria
Botswana	Guinea	Namibia	Tajikistan
Brazil	Guinea-Bissau	Nepal	Tanzania
Brunei Darussalam	Guyana	Netherlands	Thailand
Bulgaria	Haiti	New Zealand	Timor-Leste
Burkina Faso	Honduras	Nicaragua	Togo
Burundi	Hungary	Niger	Tonga
Côte d'Ivoire	Iceland	Nigeria	Trinidad and Tobago
Cabo Verde	India	North Macedonia	Tunisia
Cambodia	Indonesia	Northern Mariana Islands	Turkey
Cameroon	Iran	Norway	Turkmenistan
Canada	Iraq	Oman	Tuvalu
Central African Republic	Ireland	Pakistan	Uganda
Chad	Israel	Palau	Ukraine
Chile	Italy	Panama	United Arab Emirates
China	Jamaica	Papua New Guinea	United Kingdom
Colombia	Japan	Paraguay	United States
Comoros	Jordan	Peru	Uruguay
Congo, Democratic Republic of	Kazakhstan	Philippines	Uzbekistan
Congo, Republic of	Kenya	Poland	Vanuatu
Costa Rica	Kiribati	Portugal	Venezuela
Croatia	Korea, DPR	Qatar	Vietnam
Cuba	Korea, Republic of	Romania	Yemen
Cyprus	Kuwait	Russia	Zambia
Czechia	Kyrgyzstan	Rwanda	Zimbabwe

Chapter III

Riding the Monsoon: Geography and Iron Age Trade in the Indian Ocean

1 Introduction

Trade in the ancient world was often long, hard, and dangerous. Small wooden ships would ply the oceans with rudimentary navigational tools in the hope of landing at their destination with their produce intact. Trade costs were high, reducing the attractiveness of long-distance, trans-oceanic trade to all but the most intrepid of merchants. The Indian Ocean in the late Iron Age was an exception to this. The unique climatic elements of the seasonal monsoon made trade from the Horn of Africa to the coasts of India both feasible and attractive. A merchant setting out from one side of the Ocean could, provided they left at the correct time, travel to the other shore with relative ease.

This chapter exploits ancient textual sources to develop a database of ancient trade in the Indian Ocean and model trade in the region during the late Iron Age. Using the trade information logged in the Periplus of the Erythraean Sea (Schoff, 1912), I build a comprehensive dataset of ancient trade cataloguing products, origins, and destinations. I supplement this with wind data to calculate sailing times between the various cities along the shores of the Indian Ocean to fit a gravity model to the network. In order to perform an export analysis on the cities in the Periplus, ancient development levels are captured by archaeological data on structures and dig sites across the region.

I find that trade between cities along the shores of the Indian Ocean is well-modelled by a gravity framework, implying that long-distance trade between the continents bordering the Ocean was made feasible by the reduced travel time created by sailing conditions and monsoon winds. Further, I find that there is a significant resource endowment effect influencing trade flows throughout the region and creating a non-linear relationship between distance and trade. Cities that are very close together and have similar resource endowments trade less than cities with different endowments. This supplements the traditional technological comparative advantage of a Ricardian model with a resource-derived element of comparative advantage. I also find evidence for a relationship between the composition of exports and the level of development of cities. Cities that export more products and more diverse products tend to have higher development levels. This effect is largely driven by manufacturing products with less developed cities being more dependent on raw materials and cash-crops, such as spices and incense, implying a possible ancient version of the 'resource curse'.

This chapter makes two primary contributions. The first is to model ancient trade in an under-studied region with a gravity equation. The second is to provide insight into potentially trade-driven determinants of development in the region during the late Iron Age. In doing so, it extends the analysis of ancient trade flows into a new geographic region and adds to the discussion of driving factors of ancient trade. The majority of work on modelling trade with gravity models deals with post-eighteenth century time periods.[1] The work that has been done on ancient trade principally deals with the Mediterranean and the Middle East, rarely stretching further east or south than Mesopotamia and Egypt. This chapter extends this type of trade analysis to a new region. In contrast to Bakker et al. (2018) which find a significant role for local and regional connectedness, the peculiar conditions of the Indian Ocean allowed countries with different resource endowments to engage in sophisticated, long-distance trade with each other; leading to a non-linear relationship between distance and trade. This result is similar to that of Barjamovic et al. (2019) by finding that ancient trade in the Indian Ocean is well modelled by a gravity model but adds a non-linear element based on resource endowments. Further, the evidence contributed in this work on the link between trade and development compliments the work of Bakker et al. (2018). Whereas Bakker et al. (2018) look at the effect of connectedness on development levels, this chapter links the composition of exports and trade flows with development levels to illustrate a more direct relationship between export goods and economic activity around cities.

This work relates to the literature that documents the relationship between trade costs, communication costs, and trade flows. The unique climatic conditions of the Indian Ocean created transportation efficiencies that effectively increased the market access of cities along its shores. In this sense, the mechanisms and results of this chapter inherit much from the literature on market access. Feyrer (2009) is a good example of this with his study on changes in trade flows due to the closure of the Suez Canal, but similar work can be found in Maurer and Rauch (2019), Juhasz (2018), and Liu and Meissner (2015) among others.

The results in this chapter build on the literature around the positive effect of trade on development, including work that discusses the development effects of trade in endowment economies. The wider idea that trade links can lead to growth is a common theme throughout the development literature (see Redding and Venables, 2004; Wacziarg and

[1] Work addressing the early-modern period includes Pascali (2017) who analyses trade during the transition from sail to steam shipping.

Welch, 2008; Feyrer, 2019, for examples). I build on this tradition with the analysis of exports and development in the Indian Ocean. Another branch of literature in this area is the analysis of the 'resource curse' as outlined in Auty (1993) and Auty (2001).[2] The evidence in this chapter on the relationship between export sectors and development extends the discussion of resource endowments and growth backwards into the ancient world.

This chapter exploits historical data about ancient economies, which often require a different approach than more traditional datasets. Trade in the ancient world is frequently dealt with indirectly (as in Fluckiger et al., 2019; Broodbank, 2006) because of a lack of reliable data. Despite this, textual analysis and other techniques have proved useful in developing reliable data for research. Temin (2013) uses multiple techniques to develop useful price and trade information about Roman grain markets. Similarly, Temin (2006) and Algaze (2008) both use a combination of archaeological evidence and textual data in reconstructing trade in ancient periods. The techniques used in this chapter add to this trend by exploiting written sources for trade data.

This chapter is divided into six sections. Section 2 describes the data sources used for the analysis. A Ricardian version of a gravity model is outlined in Section 3 and Section 4 contains the results of the gravity estimation. Section 5 discusses the link between trade in the region and ancient levels of development, while Section 6 concludes.

2 Data

The data in this chapter combine archaeological, textual, and geographic information to develop a comprehensive picture of trade in the ancient Indian Ocean. The following subsections outline the different data sources and their uses in terms of the information they provide for the overall analysis.

2.1 The Periplus of the Erythraean Sea

Economic data for ancient eras are notoriously difficult to come by. Because of this, textual data have become more prominent in the literature (see Barjamovic et al., 2019; Temin, 2013; Algaze, 2008; Temin, 2006, for examples). Building on this tradition, the principal data for the gravity estimation come from a first-century document titled the Periplus of the Erythraean[3] Sea (written by an anonymous author and frequently identified by its Latin title: Periplus Maris Erythraei, PME).[4] The term 'periplus' comes from a Greek term meaning "sailing around". These types of documents were not an uncommon form of writing in ancient periods and were generally written by merchants, sailors, and other travellers to describe their journeys for the benefit of future voyagers (and also to laud their achievements).[5] This periplus was likely written by a Greek merchant in Roman

[2] Ross (1988) provides a survey of the trends in this research.

[3] The name Erythraean Sea was a Greek maritime designation for the Red Sea and Gulf of Aden that also frequently included what is now known as the Persian Gulf and the wider Indian Ocean.

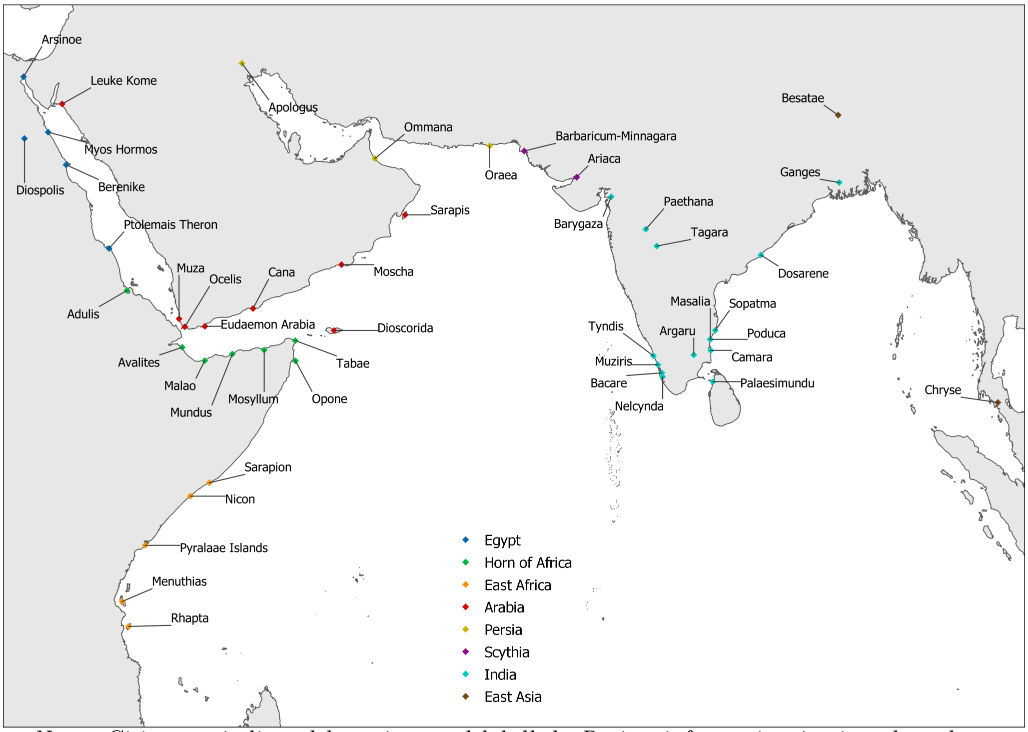
[4] The translation used for this chapter is Schoff (1912).

[5] Notable examples include the works of Hanno the Navigator, Pseudo-Skylax, and Arrian.

Egypt for the benefit of other merchants looking to ply their trade in the Indian Ocean. It follows a circular route around the coast, mentioning various cities and ports and the goods each trades. Figure 3.1 presents a map of the cities mentioned in the Periplus.

The information contained in the Periplus is remarkably complete and detailed. The source mentions trade goods, origins, destinations, product qualities, and shipment quantities. Ideally, all this information would be used to construct a comprehensive dataset of trade of a quality similar to the trade data used in modern trade analysis. Unfortunately, since its purpose was one of general description, not all of the information provided in the text is usable in this way. The information on origins, destinations, and products is the most comprehensive and usable for a detailed trade analysis, despite some gaps. The information on quality and quantity is less informative due to it being mostly subjective and relative, preventing the calculation of any value for the trade flows.

Figure 3.1. World of the Periplus



Note: Cities are indicated by points and labelled. Region information is given by colour. (Source: Author's own calculations based on Schoff, 1912).

Data Description The detail of the document provides a very comprehensive catalogue of goods exported and imported by each city mentioned and the origins and destinations of those same goods. Much of the text records a specific port of destination (origin) for each exported (imported) good:

> *Sailing through the mouth of the Gulf, after six-days' course there is another market town of Persia called Ommana. To both of these market-towns [Apologus and Ommana] large vessels are regularly sent from Barygaza, loaded with*

> *copper and sandalwood and timbers of teakwood and logs of blackwood and ebony. (PME, paragraph 36)*

This paragraph outlines specific shipments of copper and various types of timber from Barygaza to the cities of Apologus and Ommana. With this level of detail, identifying origins, destinations, and goods is a straightforward task.

The share of each broad product category in the exports and imports of the different cities is shown in Figure 3.2. The map shows a regional distribution of some exports, with different products dominating the export baskets of specific regions. This similarity in export baskets implies a regional structure to production, emphasising the potential role of resource endowments in the products offered for exports (a feature that is explored further in Section 4). This regional correlation of products is less apparent in the imports of cities. The larger variation in import goods emphasises the features of trade in the region: traders were able to travel long distances across the Indian Ocean with relative ease, making a larger variety of goods available for import throughout the network. The monsoon system contributed to the development of this more sophisticated long-distance trading network that allowed for significant expansion beyond local or regional mercantile activity.

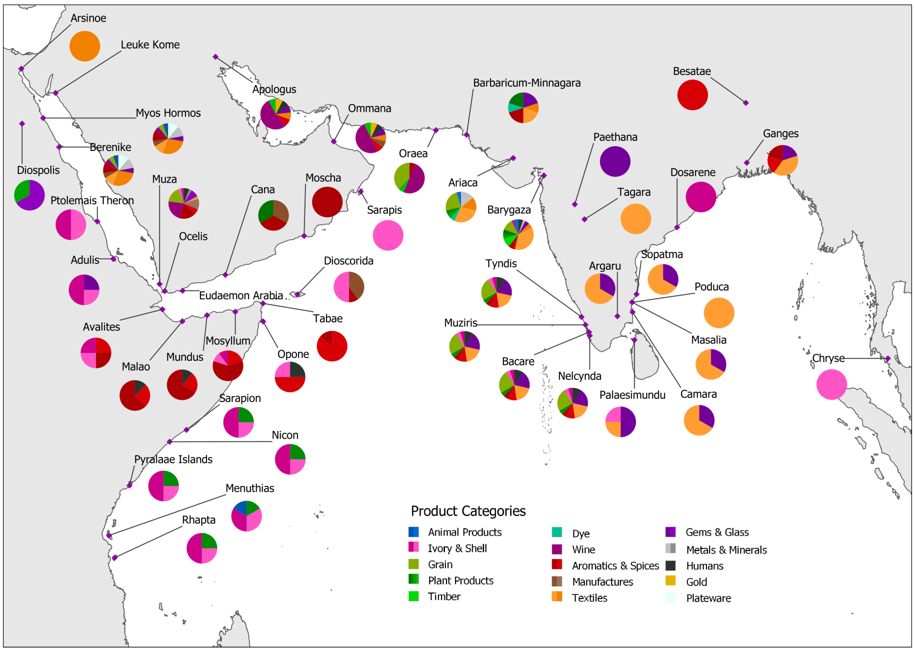
The cities mentioned in the Periplus are mostly locatable due to consistency of names across time and comprehensive archaeological work in the region. Where there is no precise location known, there is often a scholarly consensus among archaeologists about city locations. This analysis makes use of these agreed locations of the cities in the data. Figure 3.1 shows the locations and names of each city mentioned in the Periplus.

Issues Many of the passages of the Periplus follow the format above, where destination, origin, product, and other details are clearly indicated and explicitly mentioned. Despite this, there are cases in which the information is incomplete. Frequently in these cases the destination or origin of the product being traded is given as a region and not a specific city:

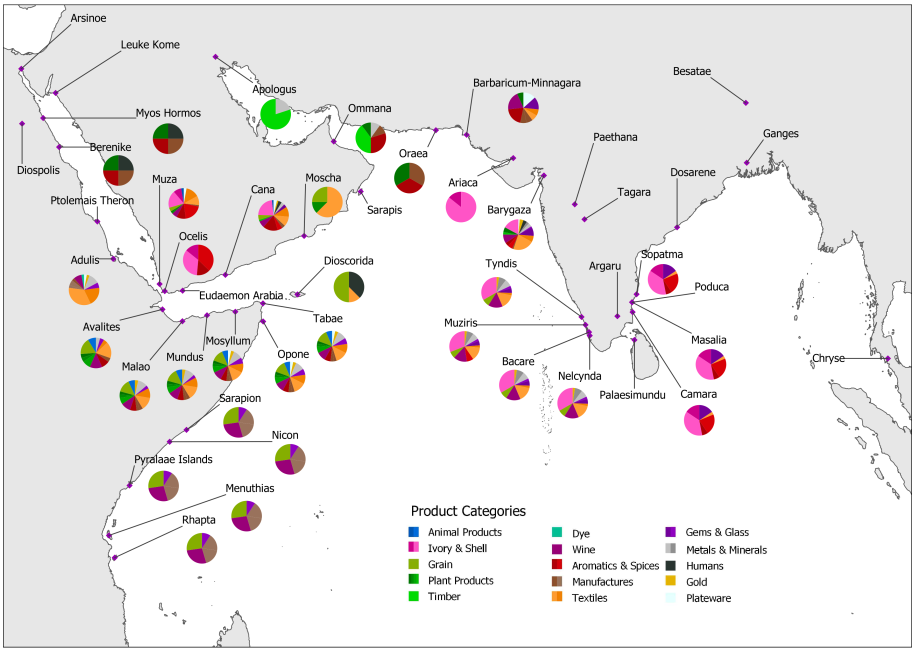
> *After Avalites there is another market-town, better than this, called Malao, distant a sail of about eight hundred stadia...There are exported from these places myrrh, a little frankincense, (that known as far-side), the harder cinnamon, duaca, Indian copal and macir, which are imported into Arabia; and slaves, but rarely. (PME, paragraph 8)*

From this paragraph, there are identified shipments of myrrh, frankincense, and other spices to Arabia, a region with no cities explicitly mentioned in this context. In cases such as this, I interpolated the relevant cities using information on the goods traded by various cities in the named region. In this way, if a city located in Arabia is identified in a different paragraph as importing myrrh or frankincense it is designated as a partner city for Avalites and Malao with those goods. Both the non-interpolated data and the interpolated data are used in the analysis.

Figure 3.2. Goods Categories as Share of Exports and Imports



(a) Export Shares



(b) Import Shares

Note: Cities are indicated by points. Pie charts display share of good categories each city exports and imports. (Source: Author's own calculations based on Schoff, 1912).

Due to the nature of the source, the terms used to describe the attributes of the trade goods are inconsistent in many places. While the name of each good is consistent across cities, the quality of the goods traded is generally rather vague and imprecise:

> *There are imported into this market-town, wine, Italian preferred, also Laodicean and Arabian; copper, tin, and lead; coral and topaz; thin clothing and inferior sorts of all kinds; bright-coloured girdles a cubit wide; storax, sweet clover, flint glass, realgar, antimony, gold and silver coin, on which there is a profit when exchanged for the money of the country; and ointment, but not very costly and not much. (PME, paragraph 49)*

As this paragraph illustrates, terms such as 'inferior' and 'not very costly' are used along with named varieties (such as Italian and Laodicean wine) or vague references to exchange rates of precious metals ('on which there is a profit'). While these passages contain significant detail about the quality or value of goods, they tend to be imprecise or relative descriptions. This prevents reliable calculations of value for the products identified in the document.

Quantities are often discussed in equally vague terms:

> *Beyond Mundus, sailing toward the east, after another two days' sail, or three, you reach Mosyllum, on a beach with a bad anchorage. There are imported here the same things already mentioned, also silver plate, a very little iron, and glass. There are shipped from the place a great quantity of cinnamon, (so that this market-town requires ships of a larger size), and fragrant gums, spices, a little tortoise shell, and mocrotu, (poorer than that of Mundus), frankincense, (the far-side), ivory and myrrh in small quantities. (PME, paragraph 10)*

As with the quality of items mentioned, phrases such as 'very little', 'small quantities', and a 'great quantity' that requires larger ships make it impossible to identify the true quantities of traded products.

These difficulties with both the quality and quantity mean that it is not possible to reliably calculate volumes or prices for goods traded by various cities, and so it is impractical to measure trade flows in value-terms. For this reason, and for the purpose of theoretical consistency that will be discussed in Section 3, trade volume is measured as the number of different goods traded. A list of different products, product categories, and the number of cities that import and export them is given in Table 3.1.

Related to the issue of incompleteness in the data is potential selective reporting of trade flows. While the Periplus is very comprehensive in its cataloguing of trade goods, trade between cities that are close together could be under-reported and trade via land routes is not recorded. In the first instance, the focus of the Periplus on larger value items and trade in differentiated goods means that trade between cities that are closer together and more likely to trade similar goods may be omitted. This could bias the results downward with the data favouring more long-distance trade. The possibility of this effect impacting the results is more fully addressed in Section 4.2.

The omission of trade via land routes is potentially more problematic with the possibility that bias is introduced through the focus on a single transportation method. While this presents a problem that is difficult to accommodate in the data, it is somewhat mitigated by the fact that trade in the ancient world was largely a marine affair, when waterways were present and accessible, due to the higher risks associated with over-land caravan trade (Gurukkal, 2016). The relative difficulty of the terrain between many of the cities in the Periplus combined with their easy access to the coast makes it likely that most trade into and out of the cities catalogued in the Periplus was conducted by sea. Despite this, the omission of land-based trade in the data remains a problem and could also lead to the results being under-estimated due to the possible concentration of such omitted trade between cities that are close together.

Table 3.1. Products of the Periplus

Product	Product Category	X	M	Product	Product Category	X	M
Clarified Butter	Animal Products	6	2	Arabian Clothing	Textiles	1	2
Horses	Animal Products	1	2	Blankets	Textiles	1	2
Mountain Tortoise	Ivory & Shell	1	1	Cloaks	Textiles	6	3
Tortoise Shell	Ivory & Shell	12	9	Coloured Sashes	Textiles	1	2
Ivory	Ivory & Shell	9	6	Girdles	Textiles	7	2
Elephant Ivory	Ivory & Shell	1	1	Local Clothing	Textiles	2	3
Rhino Ivory	Ivory & Shell	6	2	Robes	Textiles	1	1
Rice	Grain	7	7	Skin Coats	Textiles	1	1
Wheat	Grain	14	9	Thin Clothing	Textiles	1	2
Indian Copal	Plant Products	2	1	Tunics	Textiles	5	1
Aloes	Plant Products	1	12	Indian Cloth	Textiles	1	6
Coloured Lac	Plant Products	1	1	Cloth	Textiles	2	5
Macir	Plant Products	2	1	Cotton Cloth	Textiles	7	2
Palm Oil	Plant Products	5	1	Mallow Cloth	Textiles	2	2
Sesame Oil	Plant Products	7	6	Monache Cloth	Textiles	7	2
Dates	Plant Products	2	2	Muslins	Textiles	3	4
Sour Grape Juice	Plant Products	6	1	Purple Cloth	Textiles	1	2
Sugarcane Juice	Plant Products	6	2	Raw Silk	Textiles	1	5
Blackwood	Timber	2	1	Sagmatogene Cloth	Textiles	7	2
Ebony	Timber	2	1	Silk Cloth	Textiles	4	5
Sandalwood	Timber	2	1	Silk Yarn	Textiles	1	5
Teakwood	Timber	2	1	Carnelian	Gems & Glass	1	1
Purple	Dye	2	2	Coral	Gems & Glass	1	2
Wine	Wine	8	6	Diamonds	Gems & Glass	3	4
Duaca	Aromatics & Spices	2	1	Pearls	Gems & Glass	5	6
Fragrant Ointments	Aromatics & Spices	1	2	Sapphires	Gems & Glass	3	4
Frankincense	Aromatics & Spices	4	13	Transparent Stones	Gems & Glass	3	4
Myrrh	Aromatics & Spices	3	3	Flint Glass	Gems & Glass	1	1
Spikenard	Aromatics & Spices	3	4	Glass	Gems & Glass	5	1
Storax	Aromatics & Spices	1	2	Murrhine Glass	Gems & Glass	1	1
Cinnamon	Aromatics & Spices	2	3	Copper	Metals & Minerals	3	3
Malabathrum	Aromatics & Spices	4	8	Iron	Metals & Minerals	1	1
Pepper	Aromatics & Spices	3	4	Steel	Metals & Minerals	1	1
Spices	Aromatics & Spices	1	2	Tin	Metals & Minerals	1	2
Awls	Manufactures	5	1	Slaves	Humans	6	11
Daggers	Manufactures	5	1	Gold	Gold	2	2
Hatchets	Manufactures	5	1	Gold Plate	Plateware	1	2
Images	Manufactures	1	2	Silver Plate	Plateware	1	2
Lances	Manufactures	5	1				
Sewn Boats	Manufactures	1	1				
Native Produce	Manufactures	1	12				

Note: This table includes all individual products listed in the text of Schoff (1912). X indicates the number of cities exporting the product; M indicates the number of cities importing it.

Trade Variable Construction In order to construct the variable for trade used in this analysis, the number of different products a city imports or exports is summed across all product data. This is a straightforward process for bilateral trade between cities:

$$X_{ij} = \sum_{g} x_{ijg}$$
$$\text{with } x_{ijg} = \begin{cases} 1 & \text{if good } g \text{ is traded between cities } i \text{ and } j; \\ 0 & \text{otherwise.} \end{cases}$$

Where x_{ijg} is an indicator function for whether a given product g is traded between cities i and j.

The same process is used to calculate the total number of products X_i a given city i exports or imports across all partner cities:

$$X_i = \sum_{g} x_{ig}$$
$$\text{with } x_{ig} = \begin{cases} 1 & \text{if good } g \text{ is traded by city } i; \\ 0 & \text{otherwise.} \end{cases}$$

Importantly, for this measure products are not double counted. If a city trades the same item with multiple cities, it is calculated as one product making this a strict measure of the number of product varieties traded.

Table 3.2 lists the cities from the Periplus, their coordinates, the number of goods they trade, and the number of other cities they trade with. The trade network mapped by the Periplus is quite heterogeneous; the cities in various regions do not all trade with the same number of partners nor the same number of products. Many cities have only one or few trade partners. The majority of cities trade in multiple goods with only nine cities having fewer than ten products, and only three having a single product. Cities in the Horn of Africa, Arabia, Persia, and Indian are at the centre of the network with the most trade partners and trading in more products with other regions being largely peripheral to the system.

As indicated in Table 3.2, not every city in the Periplus trades with every other city mentioned. While there are no zero-trade flows explicitly recorded, these missing connections between cities could be inferred to be zeros in the trade data. This feature of the data is used in Section 4 to include zero-trade flows in the analysis.

2.2 Ancient Sea Routes

The distances between cities in the data are calculated using sailing times with respect to wind speed and direct sailing times without accounting for variations in wind speed. These travel times are calculated based on wind speed information and calculations of sailing speeds relative to the wind using the results of sea-trials of a modern reconstruction of a historical trading vessel.

Table 3.2. Cities of the Periplus

City	Region	Coordinates	Goods	Partners
Arsinoe	Egypt	29.97, 32.55	3	7
Berenike	Egypt	23.91, 35.48	22	3
Diospolis	Egypt	25.72, 32.61	3	7
Myos Hormos	Egypt	26.16, 34.24	22	3
Adulis	Horn of Africa	15.26, 39.66	32	3
Avalites	Horn of Africa	11.35, 43.47	22	6
Malao	Horn of Africa	10.44, 45.02	30	9
Mosyllum	Horn of Africa	11.17, 49.11	34	5
Mundus	Horn of Africa	10.9, 46.92	31	9
Tabae	Horn of Africa	11.82, 51.25	32	5
Opone	Horn of Africa	10.42, 51.27	30	7
Menuthias	East Africa	-6.13, 39.33	13	1
Nicon	East Africa	-1.22, 41.84	11	1
Pyralaae Islands	East Africa	-2.27, 40.9	11	1
Rhapta	East Africa	-7.85, 39.78	11	1
Sarapion	East Africa	2.03, 45.33	11	1
Cana	Arabia	14.03, 48.34	36	16
Dioscorida	Arabia	12.51, 53.92	11	10
Moscha	Arabia	17.04, 54.43	5	8
Muza	Arabia	13.32, 43.25	47	11
Ocelis	Arabia	12.77, 43.65	14	3
Sarapis	Arabia	20.47, 58.82	1	1
Apologus	Persia	30.89, 47.58	12	4
Ommana	Persia	24.34, 56.73	16	6
Oraea	Persia	25.21, 64.63	8	1
Ariaca	Scythia	23.06, 70.62	19	7
Barbaricum-Minnagara	Scythia	24.86, 67.01	22	1
Bacare	India	9.59, 76.49	34	6
Barygaza	India	21.71, 72.99	64	15
Camara	India	11.14, 79.86	13	5
Muziris	India	10.16, 76.21	34	7
Nelcynda	India	9.32, 76.54	34	6
Paethana	India	19.48, 75.38	1	1
Poduca	India	11.9, 79.82	13	5
Sopatma	India	12.53, 80.16	13	5
Tagara	India	18.32, 76.13	4	1
Tyndis	India	10.77, 75.9	34	6
Besatae	East Asia	27.33, 88.62	1	5
Thinae	East Asia	34.35, 108.72	3	5

Note: This table only includes cities that have both goods and partners listed in the text of Schoff (1912). Coordinates are in (Lat, Lon) format; Goods indicates the number of unique products the city trades (both imports and exports); Partners indicates the number of other cities a given city trades with.

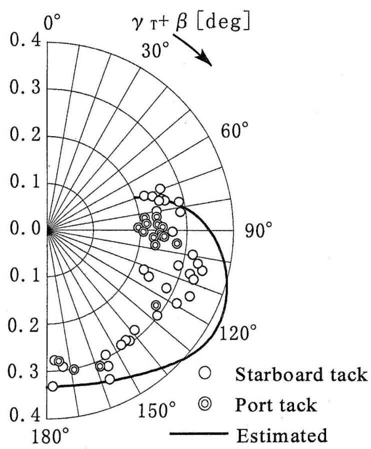
The data on wind speed and direction come from ERS-1 data on surface wind conditions from CERSAT/Laboratoire D'Océanographie Spatiale. These data provide ocean wind speed and direction information for 1 degree-by-1 degree squares of the Earth's surface. Monthly data are available over the period 1991 to 1996. The data used in this chapter are a monthly average of wind speeds over the years 1992 to 1995 for the summer monsoon (June to July) and the winter monsoon (January to February). These months are used as they tend to be peak monsoon periods and correspond with the time of year voyages would depart various ports to sail with the winds (Cobb, 2018; Gurukkal, 2016). The vector data are visualised for the Indian Ocean in Figure 3.3 and nicely illustrate the unique feature of the two monsoon seasons, with consistent south-easterly winds in the summer season compared with steady north-westerly winds in the winter season—facilitating sailing across the ocean between India and Africa/Arabia on a seasonal basis.

The speed of a sailing ship relative to the wind is based on a number of important factors, but primary among these is the angle of the ship's direction to the wind. A method of calculating this relationship is necessary in order to measure the hypothetical speed of a ship travelling with the monsoon winds shown in Figure 3.3. There are no extant data of this detail on ancient ocean-going vessels, but there do exist modern reconstructions of similar ships with accompanying tests and data. Nomoto et al. (2000) documents the sea-trials of a reconstructed Japanese *higaki kaisen*-style merchant ship, a type of ship from the sixteenth/seventeenth century and very similar in design to ancient oceanic trading vessels. The data from these sea-trials record vessel speed for given wind speeds and directions, then plots them on a polar diagram of the ship (Figure 3.4). From this information, a functional form for the speed of the ship as a proportion of wind speed for an arbitrary wind direction can be approximated with:

$$s = 0.2 - 0.129 \cos \omega$$

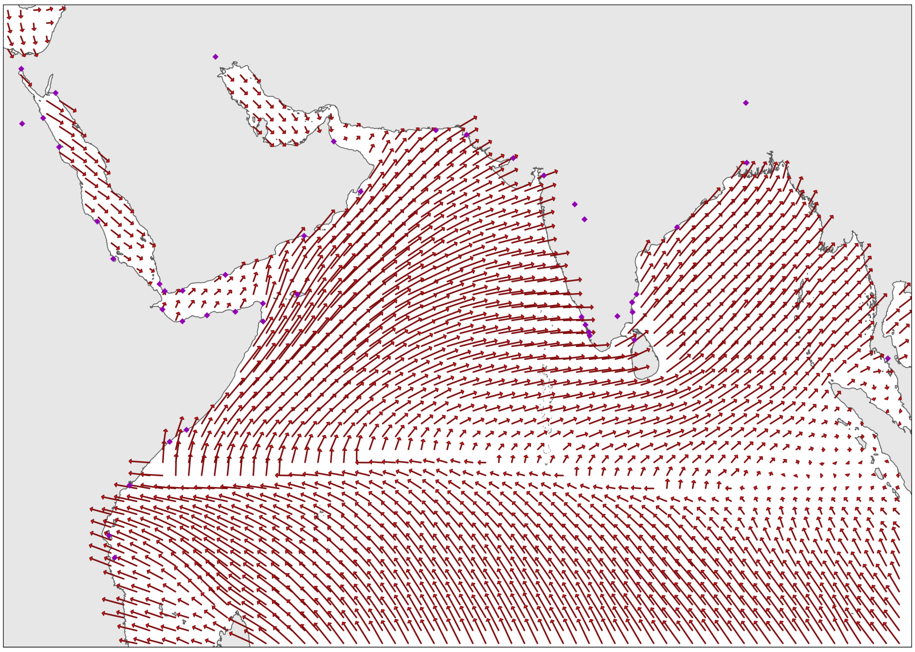
Where s is the coefficient that gives the speed of the ship as a proportion of wind speed and ω is the angle of the wind relative to the direction of travel. The first term of this formula is the average transformation of wind speed to ship speed and the second term is the amplitude around the angle of the wind. This methodology means that a ship will sail four times faster with a tailwind as opposed to a headwind. This formula also implies that a ship will travel at its fastest with a direct tailwind; a feature that, while not fully consistent with the mechanics of sailing, simplifies the process of calculating ship speeds on a grid.

Figure 3.4. Polar Diagram of the 'Naniwa-maru' (Nomoto et al, 2000)

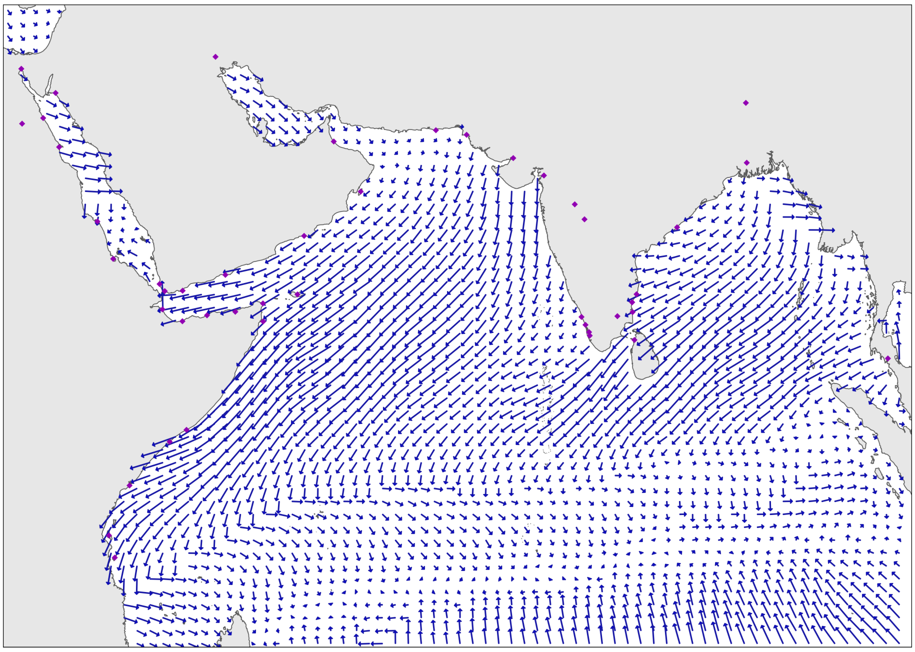


Note: Points are readings of ship speed as a proportion of wind speed at a given angle to the wind (0°). (Source: Nomoto et al., 2000).

Figure 3.3. Monsoon Winds



(a) Summer Monsoon



(b) Winter Monsoon

Note: Cities are indicated by points. Arrow direction indicates wind direction; arrow length corresponds to wind speed. (Source: Author's own calculations based on data from CERSAT/Laboratoire D'Océanographie Spatiale).

Using this formula to calculate the speed of a ship for any angle and any wind speed, a network of sailing pathways was overlaid on the wind vectors for the Indian Ocean. This created a web of pathways for a ship to travel through the 1 degree-by-1 degree grid squares of the wind vector map. The shortest and fastest routes were then calculated using GIS network analysis software to produce a series of travel times and distances along optimal sailing routes (according to wind speed and season) for all the city pairs in the Periplus. These calculated pathways are displayed in Figure 3.5. Both summer and winter routes were calculated in both directions.

Table 3.3. Travel Times Between Cities in the Periplus

Origin	Destination	Travel Time	
		Attested	Calculated
Berenike	Malao	1 Month	36.03 Days
Myos Hormos	Muza	1 Month	30.84 Days
Myos Hormos	Barygaza	>2 Months	64.70 Days
Tabae	Rhapta	<1 Month	17.99 Days

Note: Travel times are calculated according to monsoon wind data in the ideal sailing season for the direction of travel. Attested travel times come from Cobb (2018).

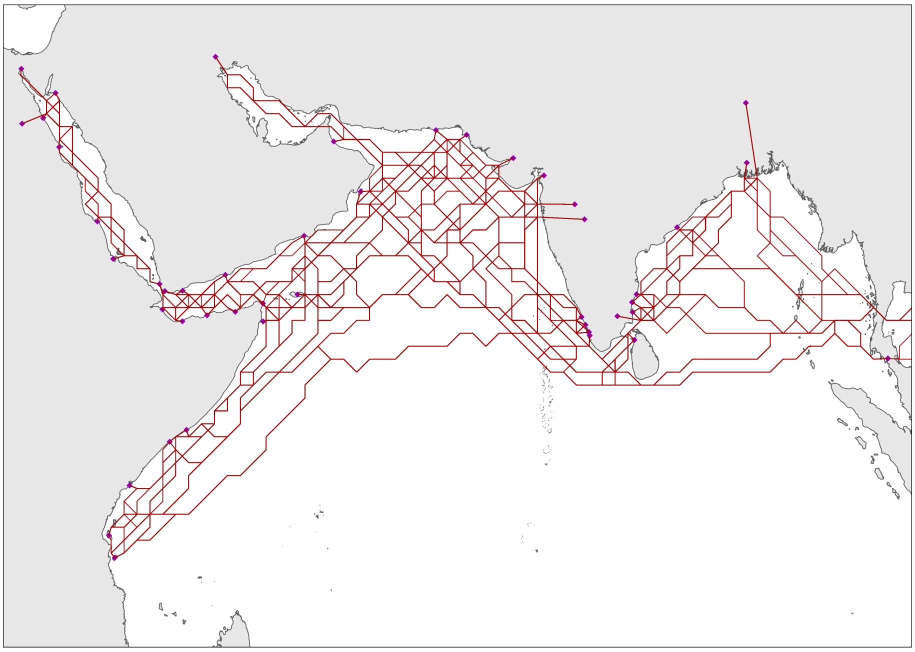
The resulting travel times are in line with attested travel times from the period and region. Table 3.3 presents selected routes and the calculated unidirectional travel times between them. The calculated distances include two types of calculations: the shortest distance in km (regardless of time), and the fastest time in days (regardless of lateral distance).

2.3 Historical Economic Activity

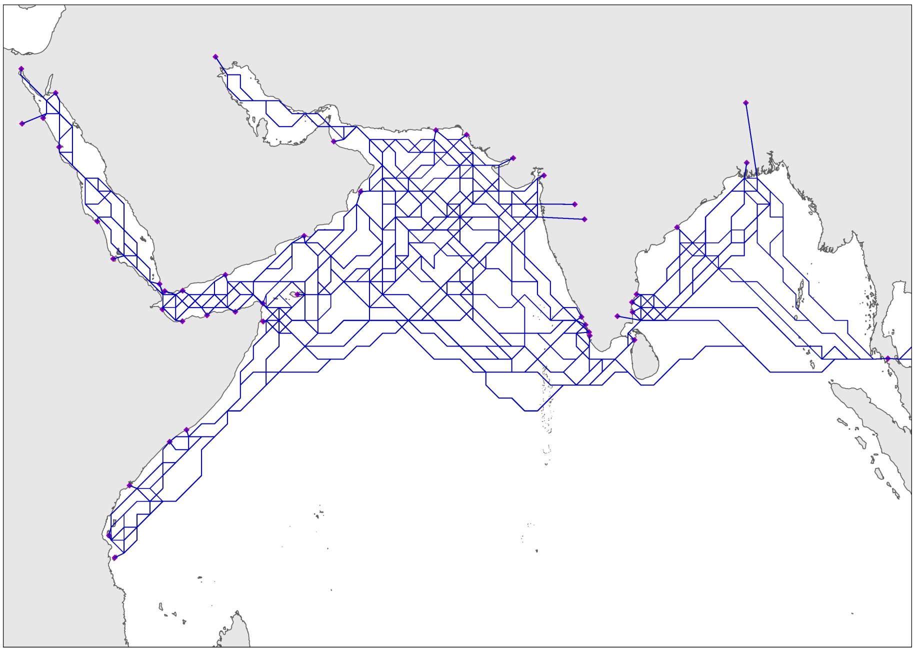
The data for historical economic activity come from archaeological sources. To proxy for historical activity, a measure of the number of built-up structures in a given location is used. This is similar to a commonly used method for determining development in modern times: the use of night-time light levels (see Elvidge et al., 2012; Bluhm and Krause, 2019, for notable examples). The concentration of human-made structures gives an indication of the amount of activity at a location at a given time.

To construct this measure, data on archaeological sites are used. The Pleiades dataset (Bagnall et al., 2016) catalogues various ‘places’ throughout ancient history. These places can be anything from provinces, temples, roads, and beyond. Each item is geolocated and given a date range for activity. This dataset originally included primarily the Mediterranean area, but its expansion to include other areas of the world (particularly the Middle East and India) makes it well suited for the purposes of this chapter. Figure 3.6 presents the location of the sites contained in the data for the region of the Periplus.

Figure 3.5. Sailing Routes Between Cities



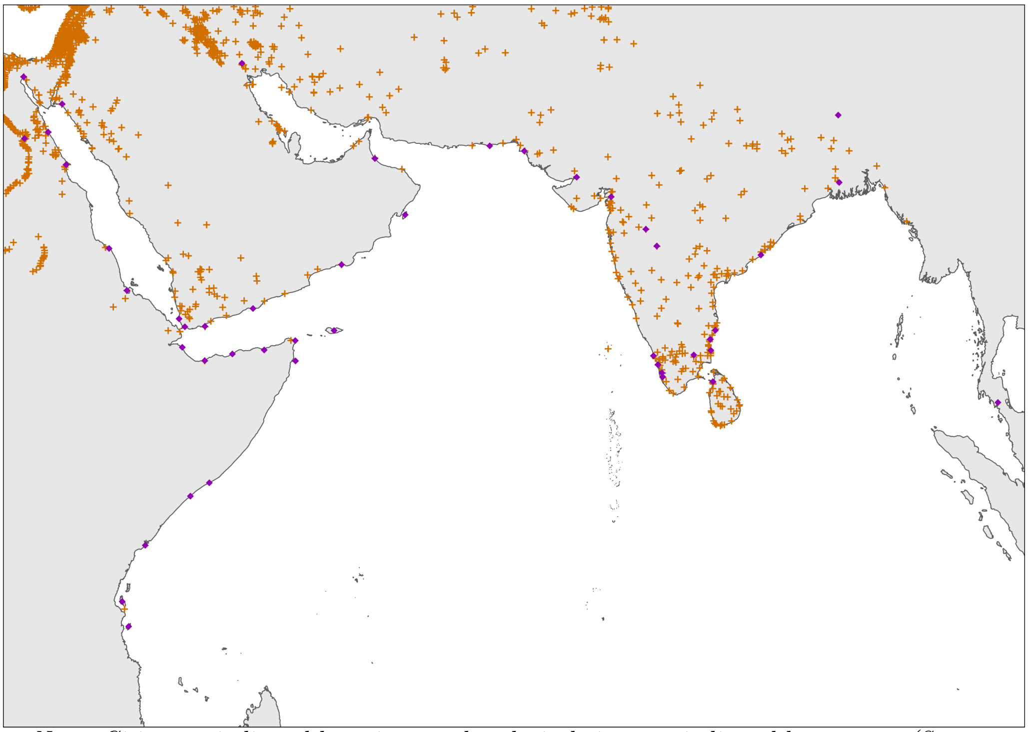
(a) Summer Routes



(b) Winter Routes

Note: Cities are indicated by points. Lines represent sailing routes between cities based on fastest sailing time. (Source: Author's own calculations).

Figure 3.6. Ancient Archaeological Sites



Note: Cities are indicated by points; archaeological sites are indicated by crosses. (Source: Author's own calculations based on Schoff, 1912; Bagnall et al., 2016).

The data used for this analysis include all human-made 'places' that existed continuously from 100 CE (around the time of the writing of the Periplus) up until 640 CE, encompassing approximately 6600 sites. The date 640 CE was chosen as it coincides with the Arab invasion and conquest of Egypt and marks the point at which Roman access to the Indian Ocean was interrupted, signifying a potentially significant disruption to the trading system of the region (McLaughlin, 2018), despite the continued functioning of the system long after (Chaudhuri, 2005). The variable for the number of sites y_i around a given city i is calculated as the sum of all sites σ within a given radius $\bar{d}_i$ of the city i:

$$y_i = \sum_{\alpha} \sigma_\alpha R_{i\alpha}$$

$$\text{with } R_{i\alpha} = \begin{cases} 1 & \text{if } d_{i\alpha} \leq \bar{d}_i; \\ 0 & \text{otherwise.} \end{cases}$$

Where R_α is an indicator function for the distance between site α and city i being less than the radius $\bar{d}_i$. These radii can be considered to correspond to a 'hinterland' view of the cities and their surroundings. The area around any given city in the ancient world is largely dependent on the prosperity of that city so that each city-hinterland area creates a core-periphery dynamic.

Table 3.4 shows the distribution of these sites across regions. There are two issues that emerge from the data on archaeological sites: one related to the summing of different sites and the other related to bias in the archaeological record. For the first issue, summing across diffferent sites captures different levels of information. The variety of different ‘places’ in the data means that summing across all sites would combine cities, towns, structures, geographical features, and political units like provinces. To address this, the type of sites used to construct this variable are limited to human-built features. While this definition includes cities, towns, and structures together, it is the most direct solution to the difficulty introduced by the second issue: a lack of data for the non-Mediterranean world.

Table 3.4. Distribution of Archaeological Sites

Region	Sites	Standard Deviation
Egypt	133.200	95.798
Horn of Africa	2.286	1.380
East Africa	0.400	0.548
Arabia	6.500	8.573
Persia	23.333	35.218
Scythia	4.500	2.121
India	19.200	9.344
All Regions	24.795	50.567

Note: Sites is the average number of sites within 200km of cities in the given region.

The second issue is related to selection bias within the data. Historically, archaeologists have been more interested in, and therefore more focused on, excavating and researching sites related to the Roman Empire or Egypt. This creates a bias within the data since the vast majority of archaeological sites are from around the Mediterranean or within the former Roman province of Egypt, with relatively few outside of these regions. While this analysis is not focused on the area around the Mediterranean, Egyptian cities are included in the data as major ports involved in trade in the Indian Ocean. This creates a problem where the number of sites in the areas around Egyptian cities is artificially large, as can be seen in Table 3.4.

In addition to this bias in the data, the fact that Egypt was a constituent part of the Roman Empire at the time introduces a significant confounding variable for the number of sites present: it is reasonable to assume that the number of sites could also be influenced by the presence of a somewhat centralised Roman administration. The other regions, though influenced by trade and economic exchange with the Roman Empire (McLaughlin, 2018), are more distant and separated from direct Roman influence meaning the number of sites is more related to local factors. In an effort to address the issue of bias in site density and the direct influence of Roman administration, I exclude Egyptian cities from the analysis of ancient development in Section 5.

3 Gravity in a Ricardian World

Trade costs in the ancient world were notoriously high; with long-distance, trans-oceanic networks exceedingly rare. In this period, distance is frequently thought of as being such a significant barrier to trade that its impact is exponential – approaching prohibitive levels of cost very rapidly. This raises the question of whether a gravity model is suitable, since if distance has a cost-prohibitive upper-bound it cannot be modelled smoothly with a standard gravity equation. While this may be a potential obstacle in other locations, the unique features of the Indian Ocean, notably the seasonal monsoon,[6] mitigate this barrier and make long-distance trade feasible (Chaudhuri, 2005). The regularity of the monsoon and the orientation of landmasses around the Indian Ocean means that regular voyages can leave any given port at a certain time of year and arrive with favourable winds at a destination on an opposite landmass. This seasonal scheduling of voyages is attested to in the Periplus and other extant sources (see Schoff, 1912, for examples) and lends evidence to the theory that predictable monsoon winds allowed for long distance trade in the region independently of advanced nautical technology.

Despite these unique features of trade in the Indian Ocean, modern gravity models that are based on monopolistic competition and demand for variety are not always well specified to deal with pre-modern trade – particularly ancient trade. In this period, production and trade of goods was largely based on technology.[7] In this context, a Ricardian world of comparative advantage predominates where specialisation is promoted through declining transport costs which facilitate trade.

In addition to a Ricardian model being more theoretically sound for ancient trade flows, the mechanics of the model also fit more precisely with the data. Due to the nature of the Periplus data described in Section 2.1 the trade variable of observation is the number of good types shipped between cities. This extensive-margin measure makes models based on the value or quantity of goods traded unworkable. With a Ricardian gravity model, the driving force of comparative advantage means that more competitive[8] cities will trade more types of products (not just more goods) and this fact can be used to fit the data to the model.

I follow Eaton and Kortum (2002) in constructing a simplified version of their gravity model of Ricardian trade.[9] In this model cities have differential access to technology. City j's technological efficiency of producing good g is given by the parameter $z_j(g)$ which is

[6]There could be debate as to the declining impact of this theoretical upper-bound as technology for ocean-going ships developed and changed in the Roman Empire (McLaughlin, 2018). Regardless, the diffusion of this technology was slow and would not have impacted non-Roman cities at the same rate, leaving the climatic factors of the Indian Ocean system a predominant force. An alternative model of trade that imposes an upper-bound on distance is discussed in Appendix B.

[7]Though endowments can also play a part, the ability of Eurasian and East African communities to exploit those endowments was determined by the spread of technology such as iron working. Section 4 explores the inclusion of endowment effects to supplement the model.

[8]In this context, competitiveness is related to both the size of the city (labour) and the distance from the destination market (transport cost).

[9]See Eaton and Kortum (2002) for a detailed derivation of the model.

the realisation of random variable Z_j, drawn independently for each g, from a city-specific probability distribution $F_j(z) = Pr[Z_j \leq z]$. City j's efficiency distribution is assumed to be Fréchet (Type II extreme value distribution):

$$F_j(z) = e^{-T_j z^{-\theta}} \tag{3.1}$$

where $T_j > 0$ is a city specific parameter and $\theta > 1$ is a global parameter capturing the variation within the distribution. A larger T_j means a higher-efficiency draw for any given good is more likely and a larger θ indicates less variability.

From this, the cost (and under perfect competition, the price in the destination) of delivering a unit of good g produced in city j to city i is

$$p_{ij}(g) = \left(\frac{c_j}{z_j(g)}\right) d_{ij} \tag{3.2}$$

where c_j is the cost of all inputs, including labour, and d_{ij} is an iceberg representation of transport costs.[10] Consumers in destination i are not limited to goods from origin j and will choose goods from destinations that minimise the prices they face. Therefore, for any given good g its price in destination i will be the lowest across all N sources:

$$p_i(g) = \min\{p_{ik}(g); k = 1, ..., N\} \tag{3.3}$$

Substituting (3.2) into the efficiency distribution (3.1) gives a distribution of prices presented to destination i by origin j:

$$G_{ij}(p) = Pr[P_{ij} \leq p] = 1 - F_j\left(\frac{c_j d_{ij}}{p}\right) = 1 - e^{-[T_j(c_j d_{ij})^{-\theta}]p^{\theta}} \tag{3.4}$$

Because the lowest price for a good in city i will be less than p unless each source j's price is greater than p, the price distribution for what city i actually buys is:

$$G_i(p) = Pr[P_i \leq p] = 1 - \prod_{j=1}^{N}[1 - G_{ij}(p)]$$

inserting the price distribution (3.4) collapses the function into something resembling $G_{ij}(p)$:

$$G_i(p) = 1 - e^{-\Phi_i p^{\theta}} \tag{3.5}$$

where

$$\Phi_i = \sum_{j=1}^{N} T_j(c_j d_{ij})^{-\theta}$$

It can then be shown that the probability that city j provides a good at the lowest price in destination i is

$$\pi_{ij} = \frac{T_j(c_j d_{ij})^{-\theta}}{\Phi_i} \tag{3.6}$$

[10] In order for 1 unit of good g to arrive in destination i, d_{ij} units must be shipped from origin j. $1 - \frac{1}{d_{ij}}$ 'melts' along the way.

which is equivalent to the contribution of city j to city i's price parameter. Helpfully, conditioning on the source has no impact on a good's price since a source with a competitive advantage exploits this by selling a wider range of goods up to the point that what it sells in a given destination i is the same as the destination's overall price distribution.[11] This means that city i's average expenditure on each good also does not vary by source, and so the proportion of goods that city i buys from city j is equivalent to the proportion of its expenditure on goods from city j:

$$\pi_{ij} = \frac{X_{ij}}{X_i} = \frac{T_j(c_j d_{ij})^{-\theta}}{\Phi_i} = \frac{T_j(c_j d_{ij})^{-\theta}}{\sum_{j=1}^{N} T_j(c_j d_{ij})^{-\theta}} \tag{3.7}$$

Departing from Eaton and Kortum (2002), this expression can then be rearranged into a equation resembling a basic gravity function in Anderson and van Wincoop (2003):

$$X_{ij} = \frac{T_j(c_j d_{ij})^{-\theta} X_i}{\sum_{j=1}^{N} T_j(c_j d_{ij})^{-\theta}} \tag{3.8}$$

using market clearing conditions, this can be turned into a standard gravity equation in the normal way:

$$X_{ij} = \frac{X_i X_j}{X_w} \left(\frac{d_{ij}}{\Phi_j \Phi_i} \right) \tag{3.9}$$

where $X_w = \sum_k X_k$ is total output and Φ_j and Φ_i serve as multi-lateral resistance terms.

Equation (3.9) is therefore a Ricardian structural gravity function and can be estimated in a similar fashion to a tradition gravity function. Following Anderson and van Wincoop (2003) and Silva and Tenreyro (2006), this model can be estimated with a pseudo-Poisson Maximum Likelihood estimator and fixed effects. With this technique, the gravity function is estimated in exponential form, allowing the number of good types to be measured in levels. The estimation equation is given by:

$$E[X_{ij}] = exp\{\beta_0 + \beta_1 ln(d_{ij}) + \mu_i + \mu_j + \epsilon_{ij}\} \tag{3.10}$$

where μ_i and μ_j represent city-specific fixed effects and standard errors are clustered at the city-pair level.

The Ricardian foundation of this model means that at its core is the concept of comparative advantage. This feature, incorporated into the price distributions, means that a city j with a higher comparative advantage (either through technology, input prices, or a shorter transport distance) will sell a wider range of goods (more product types) to city i. As the model illustrates, the proportion of goods sold to city i by city j is equivalent to the proportion of expenditure of city i on goods from city j; this is equally proportional to the range of products sold by city j to city i. This unique feature of the Ricardian model is what can be exploited in order to measure the trade between cities in terms of the number of products traded between them, rather than the value of those products as in traditional gravity models. This allows the X_{ij} term of Equation 3.10 to be given as the number of products traded between cities i and j, linking it to the analogous term from Section 2.1.

[11] See Eaton and Kortum (2002) for a detailed derivation of these properties.

4 Ancient Gravity Estimation

4.1 Linear Specification

The results of the gravity estimations using Equation 3.10 are presented in Table 3.5. The data included in these estimations are all identified trade flows, missing trade flows between cities are not included (the following subsection discusses the inclusion of these missing values as zero trade).

Table 3.5. Gravity Estimation: Distance Measures

	(1)	(2)	(3)
ln(ShortDist)	-0.269**		0.897
	(0.115)		(1.345)
ln(FastDist)		-0.270**	-1.128
		(0.113)	(1.298)
City FE	Yes	Yes	Yes
N	109	109	109
R^2	0.852	0.854	0.859

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the total number of good varieties traded; ln(ShortDist) is the log of total round-trip sailing time for the shortest calculated route without wind speed; ln(FastDist) is the log of the total round-trip sailing time for the fastest calculated route accounting for wind speed.

The first two columns of Table 3.5 compare results using the two measures of distance discussed in Section 2.2: ShortDist is the measure of the shortest geographic distance between two cities; FastDist is the measure of the fastest travel time between two cities. The coefficients on both ShortDist and FastDist are on the lower end of expected measures of the distance elasticity, but this is not uncommon under PPML estimation (Silva and Tenreyro, 2006) and could be related to the fact that the elasticity concerns the number of product types instead of explicit trade volume as in traditional gravity estimations. They are negative and significant, indicating that distance has a suppressing effect on trade—in a Ricardian world, distance reduces a city's comparative advantage in markets. In these estimations, a 1% increase in distance between two cities reduces trade between them by approximately 0.27%. The magnitudes of the two coefficients are also nearly identical, allowing for little interpretation of the suitability of each measure.

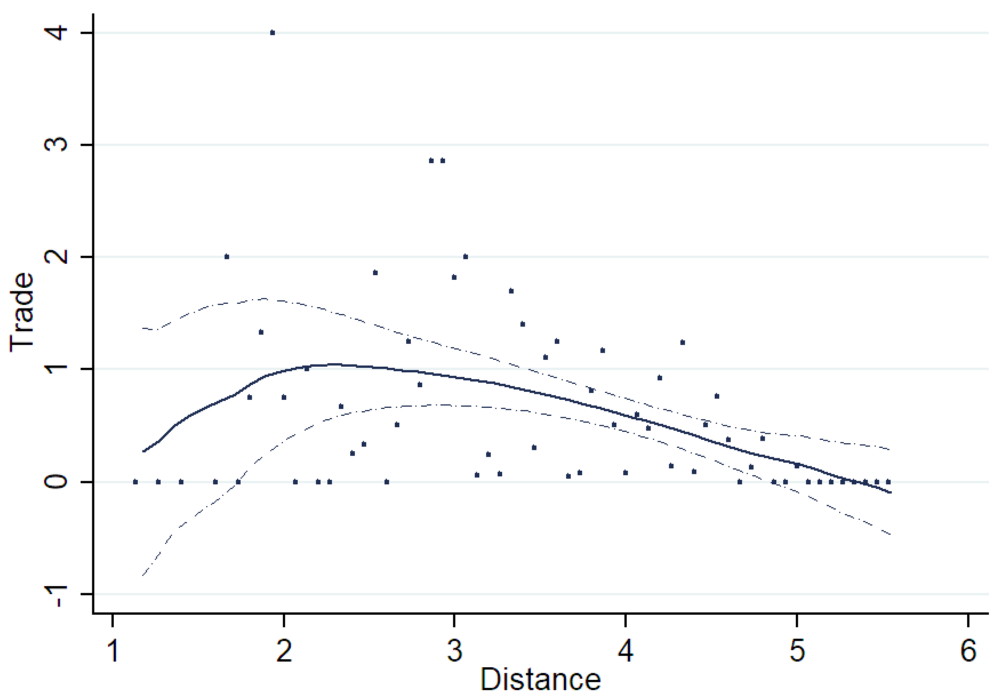
To compare the suitability of the two distance measures, I follow Pascali (2017) in creating a horse-race between the two measures. While the results are (as expected) not significant, the magnitudes and signs provide insight into the suitability of the two measures. From these results, FastDist performs better as a measure of trade distance between cities in the data, producing a strongly negative coefficient.

These results indicate that sailing time seems to be a more important factor in distance between cities than the length of the route—a result that is in line with the idea that the monsoon weather system facilitated trade in the region. The seasonal and strong winds of the monsoon created sailing efficiencies that could more than compensate for potentially longer routes.

4.2 Endowments and Non-linearity

Including city-pairs not explicitly mentioned in the Periplus as observations with zero trade removes the linearity of the relationship between distance and trade flows due to the number of cities that are close together but do not trade. As shown in Figure 3.7 the relationship takes the shape of an inverted-U due to a large number of zero-trade observations in low-distance city-pairs. Importantly, this effect is skewed towards cities that are close together, leading to two possible explanations for this effect. The first is that there could be a bias in the data, with trade between cities that are close together being systematically excluded from the Periplus because of its focus on longer-distance trade. The second is that cities with similar endowments are likely to trade less, and cities that are close together geographically are more likely to have similar endowments.

Figure 3.7. Non-linear Trade-Distance Estimates



Note: Trade is given by the total number of good varieties traded; distance is the log total round-trip sailing time for the fastest calculated route accounting for wind speed. Missing city pairs are included as zero-trade observations.

While the first explanation is plausible, it requires the vast majority of trade to be ignored regardless of the scale of such trade. It is more likely that trade would have been ignored between cities that are close together because trade between such cities was already quite low in value.

This seems to imply that a pure Ricardian model based on technology-derived comparative advantage is insufficient to fully explain the trade relationships described by the data. I therefore relax the technology-derived comparative advantage assumption of the model derived in Section 3 and incorporate factor endowments, in line with a Heckscher-Ohlin model. This allows for comparative advantage to also be derived from factor endowments.[12] Such a model is more likely to fit the data and more accurately describe trade flows.

To test this, I create two measures of proximity between cities. The first is a dummy variable indicating whether the two cities are within 14-days' sail of each other (a short return trip that could be covered easily more than once within one sailing season):

$$Neighbour_{ij} = \begin{cases} 1 & \text{if } d_{ij} \leq 14 \text{ days;} \\ 0 & \text{otherwise.} \end{cases}$$

Where d_{ij} is the sailing distance in days between city i and city j.

The second measure is an endowment index controlling for similarity in export goods offered by the cities. It is constructed as the sum across all goods categories of the difference in export baskets:

$$EI_{ij} = 1 - \sum_{n} \frac{|\lambda_{in} - \lambda_{jn}|}{L_i + L_j}$$

Where λ_{in} is the number of products within the narrow product category n (as given in Table 3.1) exported by city i and L_i is the total number of products exported by city i. Higher values of the index indicate higher levels of endowment similarity.

The results for the estimations that include these two measures are presented in Table 3.6. The first column includes just the first proximity measure with a negative and significant coefficient, indicating that there is a significant negative effect on trade when cities are close together. The second column supplements this with the endowment index. The coefficient on the index is strongly negative and significant illustrating the highly negative impact of similar endowments on trade between cities, a 1% increase in the similarity of endowments reduces the number of good varieties traded by just under 0.11%. Crucially, the coefficient on the distance measure in both estimations is much more in line with expected distance elasticities in gravity models. This seems to indicate that the lower magnitude of this coefficient in the linear analysis may have been due to these proximity and endowment effects.

While the potential for bias in the data remains due to selective reporting of trade, these results supplement the evidence for a Ricardian gravity system from the linear specification. Where the Ricardian gravity model derived for this analysis focuses on technology and distance as sources of comparative advantage, this inclusion of endowment effects

[12] Ruffin (1988) and Ruffin (1990) provide more discussion on supplementing Ricardian models with factor endowments, creating a synthesis of Ricardian and Heckscher-Ohlin trade theory.

Table 3.6. Gravity Estimation: Non-linear Distance

	(1)	(2)
ln(FastDist)	-1.098**	-1.151**
	(0.454)	(0.465)
Neighbour	-2.759**	-1.893**
	(0.719)	(0.799)
Endowment Index		-10.892**
		(1.591)
City FE	Yes	Yes
N	565	565
R^2	0.295	0.451

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the total number of good varieties traded; ln(FastDist) is the log of the total round-trip sailing time for the fastest calculated route accounting for wind speed; Neighbour is a dummy variable for cities in close proximity; Endowment Index is an index of similarity of endowments.

potentially adds another dimension to this: resource endowments. Increasing distance increases trade costs, therefore reducing trade, but cities could also be sensitive to endowment effects in that cities with similar product offer baskets are less likely to trade with each other. The size of the coefficient on the endowment effect in Table 3.6 indicates that this may be a very pronounced effect.

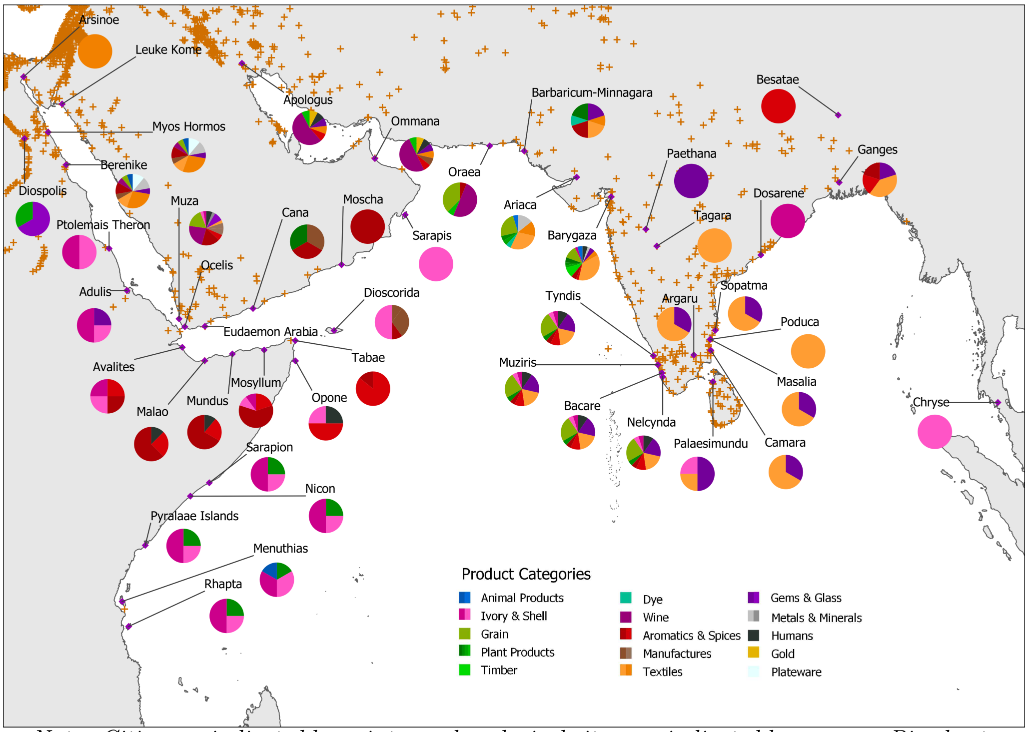
5 Trade and Development

The results of the previous section provide evidence for the Indian Ocean as a highly functioning trade network and the data used as a good representation of that activity. The impact of this trade network on the development of cities occupying the Indian Ocean littoral is the subject of this section.

5.1 Export-led Growth in the Iron Age

I use the data on exports and imports for individual cities (not city-pairs) and the data on archaeological sites from the Pleiades dataset as a proxy for ancient economic activity to test whether the trade balance of a city is related to its level of development. A frequent theme among ancient writers when discussing trade between Rome and the Indian Ocean was the threat of import dependence and the adverse economic effects of it (Cobb, 2018; McLaughlin, 2018). Generally these discussions focused on the effect of dependence

Figure 3.8. Exports and Archaeological Sites



Note: Cities are indicated by points; archaeological sites are indicated by crosses. Pie charts display share of good categories each city exports. (Source: Author's own calculations based on Schoff, 1912; Bagnall et al., 2016).

on imports from the Indian Ocean on Rome and its wealth,[13] but implicit in this is the corollary that the cities of the Indian Ocean were growing and benefiting from their exports. Figure 3.8 shows a map of the sites in the Pleiades dataset overlaid with the sectoral export shares from Figure 3.2. Cities in Egypt are excluded from the analysis for the reasons covered in Section 2.3.

For this section, the development of a city (y_i) is constructed as the number of archaeological sites within a 200km radius[14] of the city's location and is estimated as a linear combination of the effects of exports and imports using a PPML structure so that the dependent variable can be measured in levels:

$$E[y_i] = exp\{\beta_0 + \beta_1 O_i + \beta_2 Q_i + \nu_i + \epsilon_i\}$$

Where O_i is the number of export markets city i exports to and Q_i is the number of import markets. The number of export (import) markets a city trades with are calculated as the weighted sum of the number of cities a given city i exports to (imports from). The weights used are the number of goods exported to (imported from) each city. Region fixed effects are given by ν_i and correspond to the regions in Figure 3.1; ϵ_i is a normally distributed error term.

[13]Pliny complained at length about the large outflow of money to the cities and countries of the Indian Ocean and how they were becoming rich at the expense of Rome.

[14]The results of different hinterland sizes are presented in Appendix A.

Table 3.7. Ancient Development: Exports and Imports

	(1)	(2)	(3)
Export Partners	0.379**		1.091**
	(0.133)		(0.345)
Import Partners		-0.284**	-1.088**
		(0.107)	(0.432)
Region FE	Yes	Yes	Yes
N	26	30	22
R^2	0.558	0.681	0.823

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the total number of ancient place records within 200km of a city in the Periplus; Export Partners and Import Partners are logged values of the number of export or import partners weighted by the number of goods.

The results for this estimation are presented in Table 3.7. The first two columns present the results for export partners and import partners separately. These indicate that the number of export partners has a positive relationship with the level of development around the city while the number of import partners has the opposite effect. An increase in the number of export partners is linked to a 38% increase in the number of sites around a city, while a similar increase in the number of import partners reduces the number of sites by 28%. These results are reinforced in the third column which includes both export partners and import partners. Both coefficients are significant and the coefficient on export partners is strongly positive while the number of import partners is strongly negatively associated with development.

From the data, cities that have broader export markets available for their goods tend to be more developed while cities that are dependant on many import partners are under-developed. Trade in the early first century CE, specifically the number of export or import markets, seems to have influenced the number of sites present around cities during the following 500 years. This appears to indicate an effect similar to that of modern export-led growth where cities fuel their development and urbanisation through positive trade balances.

5.2 Ancient Resource Curse

The results for export-led effects on development in this section can be expanded to investigate the role of different sectors (such as manufacturing, agriculture, and luxury items like spices and incense). In order to accommodate this, two types of variables are constructed: an index of export diversity across sectors, and a set of sector-specific export variables. These capture two different elements of city export diversity. The first measures the diversity of export products while the second accounts for the concentration of those products in specific sectors that constitute different production methods (such as manufacturing, agriculture, or luxury crops).

The measure for export diversity is a modified Herfindahl index across the exported sectors of a city:

$$h_i = 1 - \sum_m \left(\frac{\lambda_{im}}{L_i} \right)^2$$

Where λ_{im} is the number of products within the sector m exported by city i and L_i is the total number of products exported by city i. This measure is constructed so that higher values indicate a higher level of sectoral diversity.

The set of sector specific export variables are calculated as simple sectoral shares:

$$S_{im} = \frac{\lambda_{im}}{L_i}$$

The share S_{im} of sector m for city i is the number of products λ_{im} in sector m exported by city i divided by the total number of products exported by city i, L_i.

These variables feed into the following estimation equation for the development of a city (y_i) given by the number of archaeological sites within 200km of the city:

$$E[y_i] = exp\{\beta_0 + \beta_1 h_i + \beta_2 S_i + \nu_i + \epsilon_i\}$$

Where h_i is the diversity index and S_i is a vector of export sector variables; ϵ_i is an error term. Regional fixed effects, ν_i, are included to control for endowment effects and regional variation in the number of archaeological sites.

The results of these estimations are presented in Table 3.8. The first three columns of the table present the diversity index and the sector-specific variables as separate models. From the first column, diversity of exports matters, with the coefficient on the diversity measure being strongly positive and significant, linking a 1% increase in the range of export sectors with a proportionate increase in levels of development. This result, while informative, could be linked to a Ricardian production effect in that more development (a larger population) could be driving increased diversification of products due to an increased comparative advantage. To add more detail to this picture, the sectoral analysis is important.

The second two columns present the sector-specific variable models. While the results in the second column are not significant, the magnitudes are suggestive of a split in impact across sectors with some contributing positively (such as manufactures) and some negatively (such as luxuries and raw materials). The third column divides the manufacturing sector into non-textile manufactures and textiles. These results are more illustrative of the possible impact of the various sectors: manufacturing has a large, positive, and significant relationship to development while luxury items (such as spices) and raw materials (such as minerals and timber) are negatively associated with the number of sites.

This seems to imply a divergence in the results for manufacturing (positive with development) and more plantation/extractive production (negative with development). Most

Table 3.8. Ancient Development: Export Diversification

	(1)	(2)	(3)	(4)	(5)
Diversity Index	1.000**			1.170*	0.556
	(0.322)			(0.608)	(0.597)
Manufactures		2.953		2.762*	
		(2.433)		(1.649)	
Manufactures (not textiles)			5.521**		5.228**
			(1.831)		(1.936)
Textiles			-3.840**		-3.154**
			(1.228)		(1.432)
Agriculture		1.851	-1.678	0.340	-1.841
		(1.533)	(1.289)	(1.603)	(1.375)
Luxuries		-0.744	-4.325**	-0.917**	-3.742**
		(0.469)	(1.290)	(0.437)	(1.564)
Raw Material		-2.404	-8.557**	-3.986	-8.233**
		(3.481)	(3.137)	(2.662)	(3.240)
Region FE	Yes	Yes	Yes	Yes	Yes
N	39	39	39	39	39
R^2	0.542	0.580	0.626	0.596	0.632

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the total number of ancient place records within 200km of a city in the Periplus; Diversity Index is a HHI-style measure of export sector diversity, a higher value is more diverse; Manufactures, Textiles, Agriculture, Luxuries, Raw Materials are the proportion of total exports that each sector represents.

of the products included in the luxuries and raw materials sectors are plantation-based cash crops (such as various spices, mining, timber, and others). These seem to relate to under-developed cities compared to those that rely more on manufacturing. The outlier here is the textile sector, which seems to be linked with lower levels of development despite being a manufacturing-related sector. This could be due to much of textile production involving many production methods similar to that of plantation-products (such as spices and incense) with much of the process dominated by harvesting and processing the raw material (frequently cotton, wool, or silk) with potentially less emphasis on weaving and dyeing.

The final two columns include the diversity index with the sector-specific variables. These results show that rather than diversity of exports, it appears that specific sectors are more important for levels of development. The coefficients display the same trends as the previous two columns in that manufactures (particularly with textiles excluded) are associated with higher levels of development while luxury goods and raw materials are negatively associated with the number of archaeological sites. This effect is analogous to the modern effect of a 'resource curse' where an over-reliance on exploiting natural resources (in this case spices, incense, and other raw materials) leads to a cycle of under-development. These results are robust to adjustments in both the cut-off year for inclusion of sites in the data and the extent of the 'hinterland' of a city. Variations in these values and the related tables are discussed in Appendix A.

These results seem to indicate that export sector shares and product diversity in trade circa 50 CE to 100 CE had significant influence on the density of archaeological sites during the subsequent 500 years to 640 CE. While causal statements in this regard are difficult due to the peculiarities of the data, this evidence is suggestive of features of the Indian Ocean trade network that are analogous to an export-led development effect and a effect similar to the modern 'resource curse'.

6 Conclusion

This chapter uses extant textual sources to model and estimate ancient trade flows in the Indian Ocean. The detail of the data allows for the construction of a comprehensive picture of economic interaction, fuelled by climatic conditions, between communities around the Indian Ocean and the way those communities were potentially shaped by that interaction. Two important results are derived from the analysis of these connections.

The first result is that the unique climatic conditions of the Indian Ocean (namely the presence of the seasonal monsoon) facilitated long-distance trade between cities along its shores. This trade is well fitted by a gravity model using sailing times between destinations based on seasonal winds. The relationship between distance and trade flows is decreasing but also non-linear, implying possible endowment effects consistent with comparative advantage being derived from both technology and resource endowments.

Second are the results for the relationship between exports and development in the region during the late Iron Age. The data and analysis presented in this chapter are suggestive of an export-driven development process with more evidence of human activity linked with increased access to export markets and specifically the influence of particular product sectors. This leads to the potential emergence of an ancient version of the 'resource curse' with cities that were more reliant on exporting the ancient equivalent of cash-crops and raw materials (such as spices, incense, and timber) being more under-developed than cities that were more reliant on manufactures.

Despite these insights, the evidence (particularly for the results related to development levels) remains suggestive. This opens up new avenues for inquiry into the potential for trade in the ancient world to influence development. Particularly inviting further investigation of the potential for and impact of an ancient 'resource curse'.

Ultimately, the insight provided by this chapter into the structure of Indian Ocean trade during the Iron Age helps to illuminate the driving forces behind economic connections in the ancient world. The results presented here are consistent with modern trade models and demonstrate the importance of long distance trade networks in the Indian Ocean during the late Classical period. Merchants and economies faced similar challenges (albeit on a different scale) to those faced by current firms and countries as the mechanisms driving the economic relationships then are analogous to those we currently face.

A Appendix: Robustness Checks

This appendix contains a number of robustness checks for the analysis presented in Section 5. The first subsection includes results for changing the date range of sites used for the export estimations in the main analysis. The second subsection covers variations in hinterland size for cities.

A.1 Third Century Decline

In the principal analysis, a cut-off of 640 CE was chosen for the presence of archaeological sites used in the data for development levels. This date was chosen as it is the date of the Arab conquest of Egypt. Despite the interruption of Roman access being a defining moment in the operation of the Indian Ocean trade network, there is evidence that Roman involvement had been in decline for several centuries prior to this date, beginning in the early third century CE.[15]

In order to address this possibility the data used for this estimation have a cut-off date of 250 CE. Table 3.9 and Table 3.10 reproduce the results of Section 5 with the new date. The implications of the results are not different from those presented in the main analysis. Specifically, the effects for exporters and importers are nearly identical; and the results for the diversity index, manufactures, luxuries, and raw materials are broadly similar with increased development linked to diversity of export sectors and manufactures in particular. Dependence on luxury goods and raw materials remain linked to under-development. This strengthens the results presented in the main analysis and provides more evidence for the possibility of both an export effect on development and a potential 'resource curse'.

Table 3.9. Third Century: Exports and Imports

	(1)	(2)	(3)
Export Partners	0.360**		0.992**
	(0.126)		(0.335)
Import Partners		-0.283**	-1.011**
		(0.099)	(0.414)
Region FE	Yes	Yes	Yes
N	26	30	22
R^2	0.572	0.697	0.811

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the total number of ancient place records within 200km of a city in the Periplus; Export Partners and Import Partners are logged values of the number of export or import partners weighted by the number of goods.

[15] Cobb (2018) and McLaughlin (2018) contain a more comprehensive discussion around the dating of this decline.

Table 3.10. Third Century: Export Diversification

	(1)	(2)	(3)	(4)	(5)
Diversity Index	0.951**			0.949	0.345
	(0.317)			(0.618)	(0.578)
Manufactures		2.947		2.728	
		(2.354)		(1.747)	
Manufactures (not textiles)			5.081**		4.893**
			(1.738)		(1.824)
Textiles			-3.515**		-3.079**
			(1.178)		(1.381)
Agriculture		2.258	-0.972	0.997	-1.064
		(1.462)	(1.165)	(1.642)	(1.235)
Luxuries		-0.809*	-4.109**	-0.970**	-3.742**
		(0.445)	(1.240)	(0.461)	(1.507)
Raw Material		-2.691	-8.237**	-3.951	-8.041**
		(3.324)	(3.126)	(2.644)	(3.217)
Region FE	Yes	Yes	Yes	Yes	Yes
N	39	39	39	39	39
R^2	0.537	0.592	0.629	0.602	0.632

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the total number of ancient place records within 200km of a city in the Periplus; Diversity Index is a HHI-style measure of export sector diversity, a higher value is more diverse; Manufactures, Textiles, Agriculture, Luxuries, Raw Materials are the proportion of total exports that each sector represents.

A.2 Variations in Hinterland Size

There is potential for the development effects analysed in Section 5 to be sensitive the the size of the 'hinterland' radius around a city. The results in this subsection include two different values for $\bar{d}_i$ in the equation from Section 2.3:

$$y_i = \sum_{\alpha} \sigma_\alpha R_\alpha$$

$$\text{with } R_\alpha = \begin{cases} 1 & \text{if } d_{i\alpha} \leq \bar{d}_i; \\ 0 & \text{otherwise.} \end{cases}$$

Tables 3.11 and 3.12 reproduce the results of Section 5 with two alternative radii: 100km and 300km. The results of Table 3.11 are broadly in line with those of the third column of Table 3.7. The coefficients have different magnitudes, particularly the ones for the 100km radius, but the signs remain the same and all are significant, supporting the trends outlined in the main analysis: increased access to export markets is correlated with higher levels of development.

Table 3.11. Alternative Hinterland: Exports and Imports

	(1) 100km	(2) 300km
Export Partners	0.952**	1.283**
	(0.401)	(0.340)
Import Partners	-0.511**	-1.089**
	(0.218)	(0.099)
Region FE	Yes	Yes
N	22	22
R^2	0.773	0.810

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the total number of ancient place records within the given radius of a city in the Periplus; Export Partners and Import Partners are logged values of the number of export or import partners weighted by the number of goods.

The results in Table 3.12 are also broadly in line with those in the main analysis. Again, the magnitudes are different across the coefficients, but the signs remain the same. Importantly, the coefficients on the key variables for manufactures, luxuries, and raw materials are significant in both specifications. This provides support for the effects outlined in the main analysis. While the coefficient on textiles is significant for a radius of 100km, it is not for the larger radius of 300km. This doesn't impact the main effects of the results, but could be related to the size of the radius in that specification—300km may be reaching the limit of what could be constructively referred to as a 'hinterland'.

Table 3.12. Alternative Hinterland: Export Diversification

	100km		300km	
	(1)	(2)	(3)	(4)
Diversity Index	0.924**	0.352	0.648**	0.741
	(0.389)	(0.625)	(0.291)	(0.596)
Manufactures		4.543**		3.054**
(not textiles)		(1.338)		(1.202)
Textiles		-2.383**		-1.031
		(1.048)		(0.866)
Agriculture		-2.215**		-0.145
		(1.143)		(0.982)
Luxuries		-2.816**		-1.723**
		(1.160)		(0.963)
Raw Material		-3.538**		-6.756**
		(2.034)		(1.750)
Region FE	Yes	Yes	Yes	Yes
N	38	38	39	39
R^2	0.639	0.744	0.626	0.596

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is the total number of ancient place records within the given radius of a city in the Periplus; Diversity Index is a HHI-style measure of export sector diversity, a higher value is more diverse; Manufactures, Textiles, Agriculture, Luxuries, Raw Materials are the proportion of total exports that each sector represents.

B Appendix: Connectedness

While this chapter and other work (such as Barjamovic et al., 2019) use a gravity model to analyse trade in the ancient world, other methods of measuring the drivers of trade have also been developed. One of these is a model based on connectedness as outlined in Bakker et al. (2018). This appendix tests the suitability of such a model for ancient trade in the Indian Ocean.

B.1 Data

The data used for the analysis of the impact of connectedness comes from purely geographic sources. In order to construct this data a map of the Indian Ocean is overlaid with a grid of cells. Each grid cell measures 0.1 degrees longitude by 0.1 degrees latitude (equivalent to approximately 11 km by 11 km). Each grid cell is then determined to either be land, sea, or coast. Land cells are those consisting entirely of land while sea cells are those entirely of sea. Coast cells are determined to be any cell with a coastline within 0.05 degrees of the centroid of the cell (this measure includes islands). Following Bakker et al. (2018), connectedness is then defined as the number of 'coastal' cells within a certain radius of a given cell—for the purpose of this appendix, a cell with a city from the Periplus in it. Table 3.13 illustrates the average number of cells in each category (coastal, land, city) for three ranges (500km, 300km, 100km) around a city.

Table 3.13. Connectedness Data

	Coast Cells	Land Cells	City Cells
500km	284.267	6951.289	5.822
	(106.417)	(1303.117)	(4.554)
300km	138.533	2502.333	3.289
	(77.441)	(470.840)	(2.312)
100km	33.000	274.733	1.622
	(23.685)	(52.588)	(1.134)

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Coast, Land, and Cities are the average number of cells within the given range of a city cell that are classified as coast cells, land cells, or cells with other cities in them. Standard deviations are given in parentheses.

B.2 Trade with Neighbours and Connectedness

The concept of connectedness as an explanation for trade relationships and development in the ancient world focuses on the idea that long-distance trade was difficult with the technology available and the opportunities for trade were limited to potential ports nearby. In this sense, the world it theorises is an extreme application of trade gravity with the connectedness model assuming that within a certain range (the limit of travel) of a given city the cost of transport is constant, outside of this range the cost approaches infinity—presenting an insurmountable barrier to trade.

From the theory of connectedness, what should impact the level of exports and imports from a city is the potential number of markets within the given range of the city. Often nearby coastline[16] is used as a proxy for this measure: the more coastline exists within the given range, the more potential for there to be markets for goods. In this way, it can be thought of as a measure of potential market access.

In pursuing this model I follow Bakker et al. (2018), incorporating many aspects of their empirical specification. While their analysis estimates the impact of connectedness on development, arguing that increased trade leads to more developed/populous locations, this appendix tests the first element of that proposed relationship directly. In doing this, I depart from their analysis by estimating the following equation:

$$E[X_i] = exp\{\beta_0 + \beta_1 c_{ri} + \psi_i + \epsilon_i\} \tag{3.11}$$

where X_i is the total trade of city i, c_{ri} is the measure of connectedness for city i (within radius r), and ψ_i is a set of geographic coordinate fixed effects. This estimation is conducted with a Poisson pseudo-maximum likelihood estimator so that trade can be measured in levels.

B.3 An Upper-limit on Navigation?

The results for the model of connectedness are presented in Table 3.14. All three specifications include coordinate controls for the centroids of the cells. Three separate distance radii are presented in this table: 500km, 300km, and 100km. No specification produced a significant result for a link between connectedness and trade levels, further the estimated coefficients themselves are close to zero with no consistent sign.

Table 3.14. Connectedness Results

	(1) 500km	(2) 300km	(3) 100km
Coast	-0.000	-0.001	0.001
	(0.001)	(0.002)	(0.008)
Coordinate FE	Yes	Yes	Yes
N	45	45	45
R^2	0.007	0.015	0.007

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable is Trade, sum of total goods traded; Coast is the number of cells with coastal geography within a given radius.

[16]Coastline is used due to the fact that sea-born trade was significantly more efficient than land-based trade for longer distances (Bakker et al., 2018).

To explore this further, the estimation is replicated for exports and imports only, includes controls for the presence of other cities within the given radius, and estimates standardised variables. Tables 3.15, 3.16, and 3.17 present the results for these alternative measures with the same ranges as before (500km, 300km, and 100km respectively). The results were not improved with these adjustments and remained insignificant.

Table 3.15. Alternative Connectedness Measures - 500km Radius

	(1) City	(2) Exports	(3) Imports	(4) Standardised
Coast	-0.001	0.001	-0.000	-0.000
	(0.002)	(0.001)	(0.001)	(0.131)
City	0.040			
	(0.059)			
Coordinate FE	Yes	Yes	Yes	Yes
N	45	45	45	45
R^2	0.016	0.025	0.029	0.007

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable (column 1) is Trade, sum of total goods traded; Dependent variable (column 2) is Exports, sum of total exports; Dependent variable (column 3) is Imports, sum of total imports; Dependant variable (column 4) is std(Trade), a standardised measure of Trade; Coast is the number of cells with coastal geography within a given radius; City indicates the number of other cities within a given radius.

Table 3.16. Alternative Connectedness Measures - 300km Radius

	(1) City	(2) Exports	(3) Imports	(4) Standardised
Coast	-0.002	0.000	-0.004	-0.115
	(0.002)	(0.002)	(0.003)	(0.160)
City	0.023			
	(0.086)			
Coordinate FE	Yes	Yes	Yes	Yes
N	45	45	45	45
R^2	0.017	0.025	0.056	0.015

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable (column 1) is Trade, sum of total goods traded; Dependent variable (column 2) is Exports, sum of total exports; Dependent variable (column 3) is Imports, sum of total imports; Dependant variable (column 4) is std(Trade), a standardised measure of Trade; Coast is the number of cells with coastal geography within a given radius; City indicates the number of other cities within a given radius.

Table 3.17. Alternative Connectedness Measures - 100km Radius

	(1) City	(2) Exports	(3) Imports	(4) Standardised
Coast	0.000 (0.009)	0.006 (0.009)	-0.004 (0.008)	0.032 (0.191)
City	0.096 (0.163)			
Coordinate FE	Yes	Yes	Yes	Yes
N	45	45	45	45
R^2	0.016	0.034	0.034	0.007

Standard errors in parentheses

* Significant at 0.10; ** Significant at 0.05

Note: Dependent variable (column 1) is Trade, sum of total goods traded; Dependent variable (column 2) is Exports, sum of total exports; Dependent variable (column 3) is Imports, sum of total imports; Dependant variable (column 4) is std(Trade), a standardised measure of Trade; Coast is the number of cells with coastal geography within a given radius; City indicates the number of other cities within a given radius.

B.4 Long Distance Trade and the Monsoon

The results above seem to indicate that a model of connectedness does not predict trade in the Indian Ocean during the Iron Age. This is not to say that the connectedness model is invalid generally, it merely demonstrates that the unique climatic conditions of the Indian Ocean reduces the efficacy of predictions based on upper-limits to navigation.

The regularity and strength of the monsoon winds in the Indian Ocean created sailing efficiencies that facilitated long-distance, inter-continental trade. In such a circumstance access to regional markets is less important for the determination of trade flows. By reducing the costs of travel between the coasts of the Indian Ocean, the monsoon winds shrunk the effective distances between cities occupying its coast, creating a more integrated and interdependent trade network over a larger distance.

Bibliography

Ades, Alberto and Hak B. Chua, "Thy Neighbor's Curse: Regional Instability and Economic Growth," *Journal of Economic Growth*, Sep 1997, *2* (3), 279–304.

Aker, Jenny C., Michael W. Klein, Stephen A. O'Connell, and Muzhe Yang, "Borders, ethnicity and trade," *Journal of Development Economics*, 2014, *107*, 1–16.

Alfani, Guido and Vincent Gourdon, "Entrepreneurs, formalization of social ties, and trustbuilding in Europe (fourteenth to twentieth centuries)," *Economic History Review*, 2012, *65* (3), 1005–1028.

Algaze, Guillermo, *Ancient Mesopotamia at the Dawn of Civilisation*, Chicago: Chicago University Press, 2008.

Anderson, James E. and Eric van Wincoop, "Gravity with Gravitas: A Solution to the Border Puzzle," *The American Economic Review*, March 2003, *93* (1), 170–192.

_ and _ , "Trade Costs," *Journal of Economic Literature*, September 2004, *42* (3), 691–751.

Arnson, Cynthia J. and I. William Zartman, *Rethinking the Economics of War: The Intersection of Need, Creed, and Greed*, Baltimore: Johns Hopkins University Press, 2005.

Austen, Ralph, *African Economic Hostory: Internal Development and External Dependency*, Oxford: James Currey, 1987.

Austen, Ralph A., *Trans-Saharan Africa in World History*, New York: Oxford University Press, 2010.

Auty, Richard M., *Sustaining Development in Mineral Economies*, London: Routledge, 1993.

_ , *Resource Abundance and Economic Development*, Oxford: Oxford University Press, 2001.

Bagnall, Roger, Richard Talbert, Tom Elliot, Lindsay Holman, Jeffrey Becker, Sarah Bond, Sean Gillies, Ryan Horne, Gabe Moss, Adam Rabinowitz,

Elizabeth Robinson, and Brian Turner, "Pleiades: A Gazetteer of Past Places," http://pleiades.stoa.org 2016.

Bah, Adama and S. Jules-Armand Tapsoba, "Civil Conflicts and Regional Economic Integration Outcomes in Africa," Technical Report, CERDI 2010.

Baier, Scott L. and Jeffrey H. Bergstrand, "Do free trade agreements actually increase members' international trade?," *Journal of International Economics*, 2007, *71*, 72–95.

Bakker, Jan David, Stephan Maurer, Jorn-Steffen Pischke, and Ferdinand Rauch, "Of Mice and Merchants: Trade and Growth in the Iron Age," Working Paper 24825, National Bureau of Economic Research July 2018.

Balassa, Bela, "Trade Creation and Trade Diversion in the European Common Market," *The Economic Journal*, 1967, *77* (305), 1–21.

Baldacchino, Godfrey, "The Contribution of Social Capital to Economic Growth: Lessons from Island Jurisdictions," *The Round Table*, 2005, *94* (378), 31–46.

Ballentine, Karen, *The Political Economy of Armed Conflict: Beyond Greed and Grievance*, Boulder: Lynne Rienner Publishers, 2003.

Barjamovic, Gojko, Thomas Chaney, Kerem Cosar, and Ali Hortacsu, "Trade, Merchants, and the Lost Cities of the Bronze Age," *Quarterly Journal of Economics*, Aug 2019, *134* (3), 1455–1503.

Bayer, Resat and Matthew C. Rupert, "Effects of Civil Wars on International Trade, 1950-92," *Journal of Peace Research*, 2004, *41* (6), 699–713.

Beestermoller, Matthias and Ferdinand Rauch, "A Dissection of Trading Capital: Trade in the Aftermath of the Fall of the Iron Curtain," *Journal of Economic History*, 2018, *78* (2), 358–393.

Blomberg, S. Brock and Gregory D. Hess, "How much does violence tax trade?," *Review of Economics and Statistics*, 2006, *88* (4), 599–612.

Bluhm, Richard and Melanie Krause, "Top Lights - Bright Cities and Their Contribution to Economic Development," Working Paper 7411, Social Science Research Network February 2019.

Bovill, Edward William, *The Golden Trade of the Moors: West African Kingdoms in the Fourteenth Century*, Princeton: Markus Wiener, 1995.

Braun, Lundy and Evelynn Hammonds, "Race, populations, and genomics: Africa as a laboratory," *Social Science & Medicine*, 2008, *67*, 1580–1588.

Broodbank, Cyprian, "The Origins and Early Development of Mediterranean Maritime Activity," *Journal of Mediterranean Archaeology*, 2006, *19* (2), 199–230.

Burchardi, Konrad B. and Tarek A. Hassan, "The Economic Impact of Social Ties: Evidence from German Reunification," *The Quarterly Journal of Economics*, 2013, *128* (3), 1219–1271.

Chae, Joon, "Trading Volume, Information Asymmetry, and Timing Information," *Journal of Finance*, 2005, *60* (1), 413–442.

Chaudhuri, K N, *Trade and Civilisation in the Indian Ocean: An Economic History from the Rise of Islam to 1750*, Cambridge: Cambridge University Press, 2005.

Chirikure, Shadreck, "Documenting Precolonial Trade in Africa," *Oxford Research Encyclopedia of African History*, 2017.

Clausing, Kimberly A., "Trade creation and trade diversion in the Canada-United States Free Trade Agreement," *Canadian Journal of Economics*, 2001, *34* (3), 677–696.

Cobb, Matthew Adam, *Rome and the Indian Ocean Trade from Augustus to the Early Third Century CE*, Leiden: Brill, 2018.

Collier, Paul and Anke Hoeffler, "The Challenge of Reducing the Global Incidence of Civil War," in Bjorn Lomborg, ed., *Global Crises, Global Solutions*, Cambridge: Cambridge University Press, 2004.

Conley, Timothy G., "Spatial Economics," in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, London: Palgrave Macmillan, 2008, pp. 1–9.

Coquery-Vidrovitch, Catherine, *The Histroy of African Cities South of the Sahara: From the Origins to Colonization*, Princeton: Markus Wiener, 2005.

Datta, Anusua and Mikhail Kouliavtsev, "NAFTA and the Realignment of Textile and Apparel Trade: Trade Creation or Trade Diversion?," *Review of International Economics*, 2009, *17* (1), 172–186.

Dunlevy, James A. and William K Hutchinson, "The Impact of Immigration on American Import Trade in the Late Nineteenth and Early Twentieth Centuries," *Journal of Economic History*, 1999, *59* (4), 1043–1062.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, Sep 2002, *70* (5), 1741–1779.

Egger, Peter H. and Farid Toubal, "Common Spoken Languages and International Trade," in Victor Ginsburgh and Shlomo Weber, eds., *The Palgrave Handbook of Economics and Language*, London: Palgrave Macmillan UK, 2016, pp. 263–289.

Eicher, Theo S., Christian Henn, and Chris Papageorgiou, "Trade creation and diversion revisited: Accounting for model uncertainty and natural trading partner effects," *Journal of Applied Econometrics*, 2012, *27* (2), 296–321.

Elvidge, C D, K E Baugh, S J Anderson, P C Sutton, and T Ghosh, "The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite data," *Social Geography*, 2012, *7* (1), 23–35.

Engel, Charles and John H. Rogers, "How Wide Is the Border?," *American Economic Review*, Dec 1996, *86* (5), 1112–1125.

Fafchamps, Marcel, "Networks, Communities and Markets in Sub-Saharan Africa: Implications for Firm Growth and Investment," *Journal of African Economies*, Sep 2001, *10* (2), 109–142.

_ , "Ethnicity and Networks in African Trade," *The B.E. Journal of Economic Analysis & Policy*, 2003, *2* (1).

Feenstra, Robert C., *Advanced International Trade: Theory and Evidence*, 2 ed., Princeton: Princeton University Press, 2004.

Felbermayr, Gabriel J., Benjamin Jung, and Farid Toubal, "Ethnic Networks, Information, and International Trade: Revisiting the Evidence," *Annals of Economics and Statistics*, 2010, (97/98), 41–70.

Feyrer, James, "Distance, Trade, and Income - The 1967 to 1975 Closing of the Suez Canal as a Natural Experiement," Working Paper 15557, NBER 2009.

_ , "Trade and Income - Exploiting Time Series in Geography," *American Economic Journal: Applied Economics*, Oct 2019, *11* (4), 1–35.

Fluckiger, Matthias, Erik Hornung, Mario Larch, Markus Lugwig, and Allard Mees, "Roman Transport Network Connectivity and Economic Integration," Working Paper 7740, Centre for Economic Studies July 2019.

Garfinkel, Michelle R., Stergios Skaperdas, and Constantinos Syropoulos, "Globalization and domestic conflict," *Journal of International Economics*, 2008, *76*, 296–308.

Geda, Alemayehu and Haile Kebret, "Regional Economic Integration in Africa: A Review of Problems and Prospects with a Case Study of COMESA," *Journal of African Economies*, 11 2007, *17* (3), 357–394.

Glick, R. and A. Taylor, "Collateral Damage: Trade Disruption and the Economic Impact of War," *Review of Economics and Statistics*, 2010, *92* (1), 102–127.

Glick, Reuven and Andrew K. Rose, "Does a currency union affect trade? The time-series evidence," *European Economic Review*, 2002, *46* (6), 1125–1151.

Goenner, Cullen F., "Simultaneity between Trade and Conflict: Endogenous Instruments of Mass Destruction," *Conflict Management and Peace Science*, 2011, *28* (5), 459–477.

Gokmen, Gunes, Wessel N. Vermeulen, and Pierre-Louis Vezina, "The imperial roots of global trade," *Journal of Economic Growth*, 2020, *25*, 87–145.

Gray, Richard and David Birmingham, *Pre-Colonial African Trade: essays on trade in Central and Eastern Africa before 1900*, Oxford: Oxford University Press, 1970.

Greif, Avner, "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *American Economic Review*, 1993, *83* (3), 525–548.

Groot, Olaf J. De, "The spillover effects of conflict on economic growth in neighbouring countries in Africa," *Defence and Peace Economics*, Apr 2010, *21* (2), 149–164.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Cultural Biases in Economic Exchange?," *The Quarterly Journal of Economics*, 2009, *124* (3), 1095–1131.

Gurukkal, Rajan, *Rethinking Classical Indo-Roman Trade: Political Economy of Eastern Mediterranean Exchange Relations*, Oxford: Oxford University Press, 2016.

Haveman, Jon and David Hummels, "Trade Creation and Trade Diversion," *Journal of Transnational Management Development*, 1998, *3* (2), 47–72.

Head, Keith and John Ries, "Immigration and Trade Creation: Econometric Evidence from Canada," *Canadian Journal of Eonomics*, 1998, *31* (1), 47–62.

_ and Thierry Mayer, "Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics*, Vol. 4 of *Handbook of International Economics*, Elsevier, 2014, pp. 131–195.

_ , _ , and John Ries, "The erosion of colonial trade linkages after independence," *Journal of International Economics*, 2010, *81* (1), 1–14.

Hegre, H., J. Oneal, and B. Russet, "Trade does promote peace: New simultaneous estimates of the reciprocal effects of trade and conflict," *Journal of Peace*, 2010, *47* (6), 763–774.

Herbst, Jeffrey, "Economic incentives, natural resources and conflict in Africa," *Journal of African Economies*, 2000, *9* (3), 270–294.

Hudson, John and Philip Jones, "International trade in quality goods: signalling problems for developing countries," *Journal of International Development*, 2003, *15* (8), 999–1013.

Hummels, David L. and Georg Schaur, "Time as a Trade Barrier," *American Economic Review*, Dec 2013, *103* (7), 2935–59.

Jedwab, Remi and Alexander Moradi, "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa," *The Review of Economics and Statistics*, 2016, *98* (2), 268–284.

Juhasz, Reka, "Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade," *American Economic Review*, 2018, *108* (11), 3339–3376.

Kahouli, Bassem and Samir Maktouf, "Trade creation and diversion effects in the Mediterranean area: Econometric analysis by gravity model," *Journal of International Trade & Economic Development*, 2015, *24* (1), 76–104.

Kang, Heejoon and Rafael Reuveny, "Exploring multi-country dynamic relations between trade and conflict," *Defence and Peace Economics*, 2001, *12*, 175–196.

King, Anne and R. M. A. van Zwanenberg, *An Economic History of Kenya and Uganda, 1800-1970*, London: Palgrave Macmillan, 1975.

Konecny, Tomas, "Expatriates and Trade," *Journal of International Migration and Integration*, 2011, *13*, 83–98.

Lewis, Roy and Yvonne Foy, "The Times Map of the Tribes, Peoples and Nations of Modern Africa," *The Times of London*, 1972.

Liu, Dan and Christopher M Meissner, "Market potential and the rise of US productivity leadership," *Journal of International Economics*, 2015, *96* (1), 72–87.

Lorenzen, Mark, "Ties, Trust, and Trade," *International Studies of Management and Organization*, 2001, *31* (4), 14–34.

Magee, Christopher S.P., "New measures of trade creation and trade diversion," *Journal of International Economics*, 2008, *75* (1), 349–362.

Maoz, Zeev and Errol A. Henderson, "The World Religion Dataset, 1945-2020: Logic, Estimates, and Trends," *International Interactions*, 2013, *39*, 265–291.

Marano, Valentina, Alvaro Cuervo-Cazurra, and Chuck C. Y. Kwok, "The Impact of Conflict Types and Location on Trade," *The International Trade Journal*, 2013, *27*, 197–224.

Marshall, Monty G., "Major Episodes of Political Violence (MEPV) and Conflict Regions, 1946-2018," http://www.systemicpeace.org/inscrdata.html 2019.

_ , **Ted Robert Gurr, and Keith Jaggers**, "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2018," http://www.systemicpeace.org/inscrdata.html 2018.

Maurer, Stephan E and Ferdinand Rauch, "Economic Geography Aspects of the Panama Canal," Working Paper 1633, Centre for Economic Performance July 2019.

Mayer, Thierry and Soledad Zignago, "Notes on CEPII's Distances Measures: The GeoDist Database," December 2011.

McLaughlin, Raoul, *The Roman Empire and the Indian Ocean*, Yorkshire: Pen and Sword Books Ltd, 2018.

Melitz, Jacques, "Language and foreign trade," *European Economic Review*, 2008, *52* (4), 667–699.

_ **and Farid Toubal**, "Native language, spoken language, translation and trade," *Journal of International Economics*, 2014, *93* (2), 351–363.

Michalopoulos, Stelios and Elias Papaioannou, "The Long-Run Effects of the Scramble for Africa," *American Economic Review*, July 2016, *106* (7), 1802–48.

Mirza, Daniel and Thierry Verdier, "International trade, security and transnational terrorism: Theory and a survey of empirics," *Journal of Comparative Economics*, 2008, *36*, 179–194.

Mitchener, Kris James and Marc Weidenmier, "Trade and Empire," *The Economic Journal*, Oct 2008, *118* (533), 1805–1834.

Murdoch, James C. and Todd Sandler, "Economic Growth, Civil Wars, and Spatial Spillovers," *Journal of Conflict Resolution*, Feb 2002, *46* (1), 91–110.

_ **and** _ , "Civil Wars and Economic Growth: Spatial Dispersion," *American Journal of Political Science*, Jan 2004, *48* (1), 138–151.

Murdock, George Peter, *Africa: Its Peoples and Their Culture History*, New York: McGraw-Hill, 1959.

Nitsch, Volker and Dieter Schumacher, "Terrorism and international trade: an empirical investigation," *European Journal of Political Economy*, 2004, *20*, 423–433.

Nomoto, Kensaku, Yutaka Masuyama, and Akira Sakurai, "Sailing Performance of 'Naniwa-maru', Reconstruction of a Japanese Sailing Trader in Early 1800s," *Journal of the Kansai Society of Naval Architects, Japan*, 2000, (234), 115–124.

Nunn, Nathan, "The long-term effects of africa's slave trades," *The Quarterly Journal of Economics*, 2008, *123* (1), 139–176.

_ **and Diego Puga**, "Ruggedness: The blessing of bad geography in Africa," *The Review of Economics and Statistics*, 2012, *94* (1), 20–36.

Ogot, Bethwell A., *Zamani : a survey of East African history*, Nairobi: East African Publishing House, 1974.

Oneal, J. and B. Russet, "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950-1985," *International Studies Quarterly*, 1997, *41* (2), 267–294.

_ **and** _ , "Assessing the Liberal Peace with Alternative Specifications: Trade Still Reduces Conflict," *Journal of Peace Research*, 1999, *36* (4), 423–442.

_ **and** _ , "The Kantian Peace: The Pacific Benefits of Democracy, Interdependence, and International Organizations, 1885-1992," *World Politics*, 1999, *52* (1), 1–37.

Panagariya, Arvind, "Preferential Trade Liberalization: The Traditional Theory and New Developments," *Journal of Economic Literature*, 2000, *38* (2), 287–331.

Pascali, Luigi, "The Wind of Change: Maritime Technology, Trade, and Economic Development," *American Economic Review*, 2017, *107* (9), 2821–2854.

Portes, Richard and Helene Rey, "The determinants of cross-border equity flows," *Journal of International Economics*, 2005, *65*, 269–296.

_ **,** _ **, and Yonghyup Oh**, "Information and capital flows: The determinants of transactions in financial assets," *European Economic Review*, 2001, *45* (4-6), 783–796.

Qureshi, Mahvash Saeed, "Trade and thy neighbor's war," *Journal of Development Economics*, 2013, *105*, 178–195.

Rapoport, Hillel and Frederic Docquier, "Chapter 17 - The Economics of Migrants Remittances," in Serge-Christophe Kolm and Jean Mercier Ythier, eds., *Handbook of the Economics of Giving, Altruism and Reciprocity*, Vol. 2 of *Handbook of the Economics of Giving, Altruism and Reciprocity*, Elsevier, 2006, pp. 1135–1198.

Rauch, James E., "Business and Social Networks in International Trade," *Journal of Economic Literature*, 2001, *39* (4), 1177–1203.

_ **and Vitor Trindade**, "Ethnic Chinese Networks in International Trade," *The Review of Economics and Statistics*, 2002, *84* (1), 116–130.

Redding, Stephen and Anthony J Venables, "Economic geography and international inequality," *Journal of International Economics*, 2004, *62*, 53–82.

Robinson, Amanda Lea, "Internal Borders: Ethnic-Based Market Segmentation in Malawi," *World Development*, 2016, *87*, 371–384.

Rose, Andrew K., "Do We Really Know That the WTO Increases Trade?," *American Economic Review*, March 2004, *94* (1), 98–114.

Ross, Michael L., "What Have We Learned about the Resource Curse?," *Annual Review of Political Science*, 1988, *18*, 239–259.

Roy, Devesh, Abdul Munasib, and Xing Chen, "Social trust and international trade: the interplay between social trust and formal finance," *Review of World Economics*, 2014, *150*, 693–714.

Ruffin, Roy J., "The Missing Link: The Ricardian Approach to the Factor Endowment Theory of Trade," *American Economic Review*, 1988, *78* (4), 759–772.

_ , "The Ricardian Factor Endowment Theory of International Trade," *International Economic Journal*, 1990, *4* (4), 1–19.

Schoff, Wilfred H., *The Periplus of the Erythraean Sea*, New York: Longmans, Green, and Co., 1912.

Selmier II, W. Travis and Chang Hoon Oh, "The Power of Major Trade Languages in Trade and Foreign Direct Investment," *Review of International Political Economy*, 2013, *20* (3), 486–514.

Silva, J. M. C. Santos and Silvana Tenreyro, "The Log of Gravity," *The Review of Economics and Statistics*, 2006, *88* (4), 641–658.

Spring, Eva and Volker Grossmann, "Does bilateral trust across countries really affect international trade and factor mobility?," *Empirical Economics*, 2016, *50*, 103–136.

Stinnett, Douglas M., Jaroslav Tir, Philip Schafer, Paul F. Diehl, and Charles Gochman, "The Correlates of War Project Direct Contiguity Data, Version 3," *Conflict Management and Peace Science*, 2002, *19* (2), 93–112.

Temin, Peter, "Mediterranean Trade in Biblical Times," in Ronald Findlay, Rolf Henriksson, Hakan Lindgren, and Mats Lundahl, eds., *Eli Heckscher, International Trade, and Economic History*, Cambridge: MIT Press, 2006, pp. 141–156.

_ , *The Roman Market Economy*, Princeton: Princeton University Press, 2013.

Viner, Jacob, *The Customs Union Issue*, New York: Carnegie Endowment for International Peace, 1950.

Wacziarg, Romain and Karen Horn Welch, "Trade Liberalization and Growth: New Evidence," *World Bank Economic Review*, 2008, *22* (2), 187–231.

Wagner, Don, Keith Head, and John Ries, "Immigration and the Trade of Provinces," *Scottish Journal of Political Economy*, 2002, *39* (5), 507–525.

Windmeijer, F. A. G. and J. M. C. Santos Silva, "Endogeneity in Count Data Models: An Application to Demand for Health Care," *Journal of Applied Econometrics*, 1997, *12*, 281–294.

Woolcock, Michael and Deepa Narayan, "Social Capital: Implications for Development Theory, Research, and Policy," *The World Bank Research Observer*, Aug 2000, *15* (2), 225–249.

Yu, Shu, Sjoerd Beugelsdijk, and Jakob de Haan, "Trade, trust and the rule of law," *European Journal of Political Economy*, 2015, *37*, 102–115.